

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 15.1684 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MAARVLLLTALTLCHGFNLD.....FKRQYKDMSEGPPGAEPQ 1153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	1 RWHU1B	cell surface glyco
2	4545	76.3	1153	2 S00551	leukocyte surface
3	3513	59.0	1163	1 RWHU1C	cell surface glyco
4	1559	26.2	1170	2 S03308	cell surface glyco
5	1549	26.0	1163	2 I56126	lymphocyte fuction
6	1163	19.5	1179	2 A53213	integrin alpha-E c
7	1108.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1099	18.5	1170	2 I45914	integrin alpha 2 s
9	1089	18.3	1178	2 S44142	VLA-2 protein homo
10	1087.5	18.3	1181	2 A33998	integrin alpha-2 c
11	1082	18.2	1180	2 A35854	integrin alpha-1 c
12	667	11.2	1039	2 A41131	lymphocyte-Peyer's
13	638	10.7	1038	2 S06046	integrin alpha-4 c
14	630	10.6	1035	2 I58409	integrin alpha-9 c
15	614.5	10.3	1041	2 T31437	integrin alpha cha
16	579.5	9.7	1054	2 JC7294	alphan integrin -
17	572.5	9.6	1051	2 A35761	cell surface glyco
18	567.5	9.5	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.3	1053	2 A34250	integrin alpha-5 c
20	550	9.2	1034	2 A36108	integrin alpha-V c
21	535	9.0	1044	2 T10050	integrin alpha-v c
22	532	8.9	1049	2 A27079	fibronectin recept
23	532	8.9	1073	2 B36429	integrin alpha-6 c
24	530.5	8.9	1072	2 A38457	integrin alpha-6 c
25	529.5	8.9	1051	2 A40021	integrin alpha-3 c
26	526	8.8	1048	2 A27421	integrin alpha-5 c
27	525.5	8.8	1091	2 A41543	integrin alpha-6 c
28	517	8.7	1044	2 T16516	integrin alpha-8 c
29	505	8.5	1394	2 A29637	position-specific

30 496.5 8.3 1146 2 S40311 integrin - fruit f
31 495.5 8.3 1039 2 A34269 integrin alpha-2b
32 492.5 8.3 1037 2 A60163 glycoprotein I1b -
33 490 8.2 1137 2 JCS950 integrin alpha-7 c
34 488 8.2 1135 2 I61186 alpha-7 integrin -
35 486 8.2 126 2 B30892 leukocyte adhesion
36 474.5 8.0 1226 2 S44824 P54F2.1 protein -
37 469.5 7.9 1106 2 S38783 integrin alpha cha
38 454.5 7.6 1139 2 S28277 hypothetical prote
39 454 7.6 1045 2 S60571 integrin alpha v c
40 430.5 7.2 1115 2 T09433 integrin alpha cha
41 423.5 7.1 1115 2 T09403 integrin alpha cha
42 391 6.6 764 2 I36916 glycoprotein I1b -
43 309 5.2 1086 2 T18523 integrin alpha cha
44 304.5 5.1 272 2 A55348 integrin alpha-1 -
45 299 5.0 604 2 I36917 glycoprotein I1b -

ALIGNMENTS

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <OR>

A:Cross-references: GB:J03925; NID:gl87284; PIDN:AA59544.1; PID:g307148

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AA5A

A:Note: the authors translated the codon TAC for residue 1129 as Thr

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally r

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FLE>
A:Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
A:Note: sequence extracted from NCBI backbone (NCBIP:121963)
R:Pietre, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PIE>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: I52567; MUID:92144986; PMID:1346576
A:Accession: I52567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:g180184; PIDN:AAA1960.1; PID:g553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A:Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VMA2>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCCHGFNLDENAMTFQENARGFGSVVOLQSGVVVGAQPIVAANQR 60
Db 1 MALRVLLLTALTLCCHGFNLDENAMTFQENARGFGSVVOLQSGVVVGAQPIVAANQR 60

Qy 61 GSYLQCDYSTGSCPIRLQVPVEAVNMSGLSLAATSPQLLACGPTVHQTCSNTYVK 120
Db 61 GSYLQCDYSTGSCPIRLQVPVEAVNMSGLSLAATSPQLLACGPTVHQTCSNTYVK 120

Qy 121 GLCFLFGSNLRQOPKPFALRCQPEDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPKPFALRCQPEDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQL 180

Qy 181 KSKTFLSLMQYSEERIHFTTFKFNQNNPNSRLVKPITOLLGRTHATGIRKVVRELFN 240
Db 181 KSKTFLSLMQYSEERIHFTTFKFNQNNPNSRLVKPITOLLGRTHATGIRKVVRELFN 240

Qy 241 ITNGARKNAPKILVITDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAPKILVITDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300

Qy 301 NTIASKPPRDHRVFNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
Db 301 NTIASKPPRDHRVFNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360

Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLYGAAAILLRNVQSILV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLYGAAAILLRNVQSILV 420

Qy 421 LGAPRYOHITGLVAMERQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Db 421 LGAPRYOHITGLVAMERQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480

Qy 481 HYTEOTRGQGVSVCPPLRGQARMCDAVLYGEOQGPWGRFGAALTVLGDNVNGDKLTQVDA 540
Db 481 HYTEOTRGQGVSVCPPLRGQARMCDAVLYGEOQGPWGRFGAALTVLGDNVNGDKLTQVDA 540

Qy 541 IGAPGEENRGAAYLFFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEENRGAAYLFFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600

Qy 601 DLTVGAQGHVLLRSQPLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660

Qy 661 STRDLREGQIQSVVTVYDLALDSGRPHGRAVENETKNSTRROTQVLGLTQTCTETLKLQLP 720
Db 661 STRDLREGQIQSVVTVYDLALDSGRPHGRAVENETKNSTRROTQVLGLTQTCTETLKLQLP 720

Qy 721 NCIEDPVPSPVLRLNFSLVGTPLSAFGLNLRPVLAEADAQRLFTALPFEKNCNDNICQDD 780
Db 721 NCIEDPVPSPVLRLNFSLVGTPLSAFGLNLRPVLAEADAQRLFTALPFEKNCNDNICQDD 780

Qy 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPLDLVSRKYSTLQNRK 840
Db 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPLDLVSRKYSTLQNRK 840

Qy 841 QRSWLACESASSTEVSGALKSTCSINHPIPEPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACESASSTEVSGALKSTCSINHPIPEPENSEVTFNITFDVDSKASLGNKLLKA 900

Qy 901 NVTSENMPRTNKTFFQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHOYQVSN 960
Db 901 NVTSENMPRTNKTFFQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHOYQVSN 960

Qy 961 LGORSLSPLSLVPLVRLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV 1020
Db 961 LGORSLSPLSLVPLVRLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV 1020

Qy 1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
Db 1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080

Qy 1081 TLLPGQAFVRVSRQETKVEPFPVNPPLIIVGSSVGLLLALITAAALYKLGFFKRYQKD 1140
Db 1081 TLLPGQAFVRVSRQETKVEPFPVNPPLIIVGSSVGLLLALITAAALYKLGFFKRYQKD 1140

Qy 1141 MMSEGPPGAEAPQ 1153
Db 1141 MMSEGPPGAEAPQ 1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C;Accession: S00551; I59078
R;Pytel, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551


```

631 EVARNVFECNDQVVGKKEAG--EVRVCLHVQ-KSTRDLRREQIQSVVTVYDLALDSGRPH 687
652 KYNIQKQNC--MEGKETVCINATVCFEVLKSKEDTIYEADLQ----YRVTLDSLRLQI 704
688 SRAVNET-----KNSTRRTOTVGLGTQTCETLKLQLPNCIEDVSPVLRLNFSVLGT 741
705 SRSPFSGTQERKVQRNITVRKSEC-----TKHSFVWLKDKHDFQSVR---ITLDFNLT-D 755
742 PLISAFGNLRPVLAEADQRLFTALPFPEKNCNGNDNICDDLSITFSFMSLDCLVGGPRE- 800
756 PENG-----PVLDDSLPNSVHEYIIPFAKDCGKKEKICISDLHLVATTEKDLLIVRSQNDK 810
801 FNVTVTVRNGDSDSVRTQVTFPFDLDSYRKVKYSTLONORSQSRWRLACESASSTEVSGAL 860
811 FNVSLTVKNTKDSAYNTRTIVHSPNLVFSGIEAIQKD-----SCESN----- 853
861 KSTSCSINHPIFPENSEVTENITDPOVDSKASIGN-KLLLKANVTSENMPRTNKTEFOLE 919
854 HMITCKGVPPFRRGEMVTFKLFQNTSYLMENVTIYLSATSDSEEPETLSDNVNVNIS 913
920 LPVKYAVYVMTSHGVSTKYLNTFASENTSRVMQHQYQVSN-----LGORS-----L 966
914 IPVKEVGLQFYS-SASEVHSIAANETVPEVINSTEDIGNEINFYLIRKSGSPMPPEL 972
967 PISLVF-----LVPVRLNQTVIWDPRQVTFSENLSSTCHTKYE-----RLPS 1007
973 KLSISFPNNMTSGYPVLPTGLSS-----SENANCRPHIFEDPFPSINSCKMKT 1021
1008 HSDFLAELRKAPVNCSTAVCORIOCDIPFGIOE-----EFNATLK 1049
1022 STD---HLKRGFTLDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSNLNTR 1078
1050 GNLSFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGCAFVRSOTETKVEPEVFNPLPL 1109
1079 GEL-----RENASLVLSN-----QKRELAIQISKDGLRQVPL 1114
1110 --IVSGVGLLLALITAAALYKLGFFKRYQKMMSE 1144
1115 WILLISAPAGLLLLMLLILALWKIGFFKRPKLKKMEK 1151

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit
A:Reference number: A54402; MUID:94193647; PMID:7511592
A:Accession: I45914
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <K>
A:Cross-references: GB:L125886; NID:9439695; PIDN:AAB59255.1; PID:9439696
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F:161-336/Domain: von Willebrand factor type A repeat homology <WA2>

Query Match 18.5%; Score 1099; DB 2; Length 1170;
Best Local Similarity 27.8%; Pred. No. 1.3e-67;
Matches 340; Conservative 215; Mismatches 498; Indels 168; Gaps 47;

QY 11 LTLCHGFNLDTENATFQ-ENARGFGQSVVL---QGSRVVVGAPQEIIVAAQRGSLVQC 66
DB 13 LNCVCAYNVGLPKAIFSGPSSEQGYAVQQPINPKGNWLLVGSWPSGFPKRNMGDVVKC 72
QY 67 --DYSGSCSEPIRLQ-----VPEAVNMSLGLSLAATSPQLLACGPTVHOTCSENT 117
DB 73 PVDLSITTCUKNLQTSMTSMNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLMAQQCGSQY 132
QY 118 YVKGLCFLFGSNLRLOPQKFFPEALRGCPQEDSDIAFLIDGSGSIIIPHOPRRNKFEVSTVM 177

```

Db 133 YTTGVCSDVDF-QURTSFAPAVOTCP-SFIDVVVVCDSESNIIYPWD--AVKNLEKXV 188
QY 178 EOLK--KSKTSLFMOYSEFRHFTFKFONNPNRSLVKPITOLL-----GRTHATGI 231
Db 189 QGLDIGPTKTQGLIOYANPRVFNLTFSKD---EMIKATSTQFYGGDLTWTFFAI 245
QY 232 RKVRELFNITNGARKNAFKILVITDGBKFGDPLGYEDVPEADREGVIRVIGV----287
Db 246 QYARDTAYSTAAGRGPCATKVMVVVTDGESH-DGSKLKAVIDQCNDILRFGIAVLGYL 304
QY 288 -GDARFSEKSRQELNITIASPPRDHVFQVNNFEALKTIONQREKIFAIEGTQTCSSSF 346
Db 305 NRNALDTKNLKEIKAINSIPTERHFNVSDEADLEKAGTIGEQIFSIEGTQVQ-GDNF 363
QY 347 EHEMSQEGFSAAT--SNGPLSTVGSYDWAGVFLYTSKESSTFINMT--RVDSDMN-D 401
Db 364 QHEMSQVGSFAEYSPONNIMLGAAGYDWSGTIVQKTHGLHIFSKQAFQILODRNHS 423
QY 402 AYLGYAAAAILNRVQSLVGLAPRYQHIGLVAMFRONTQWGESNANV-----KGTQIGA 455
Db 424 SYLGSVASISTGNSVHFVAGAPRANYTGQVLYSVN-----ENGNTVTIQSGRQDQGS 478
QY 456 YFGASLCSVDVDSNGSTDLVLICAPHYEQTR--GGQVSVCPPLPRQARWQCDVLYGE 513
Db 479 YFGSVLCAVDVWVKDITDVLVAGAPMYMNDLKEEGRVYLFITKG-ILNWH--QPLEG 535
QY 514 QGQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGIGPSHSQRTA 573
Db 536 NGLNARFGSAIALSDINMDGNDVIVGSPLENQNGAVIYNGHEGM-IRLYSQKIL 594
QY 574 GS--KLSPRIQYFGOSLGGQDITMDGLVLTGAQGHVLLRLQRSQVLRVKAIMEFNPRE 631
Db 595 GSDRAFSSHLQYFGSLDYGDLNGDSITDVSVGAFGQVQVQLMSQSIADVSVDASPTPK 654
QY 632 VARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVTDLALD-----SGRH 687
Db 655 I--TLNKNABE-----KLKCF---SAKFRPTNQNQVAIVNITIDEDQFSRVI 701
QY 688 SRAVFNETKSTRTOVLGLTQCE--TLKQLPNCIEDPVPSPVLRNLNFSL--VGTP 743
Db 702 SRGLFKENNERCLOKTMVISOQRCEYIIHIEPS---DIISPLNLCMNISLENFGT--756
QY 744 SAFGNLRPLVAEDAORLFTALPFPEKNCNDNITCDDLSITF---SPMSLCLVVGGR 799
Db 757 -----NPALAESETVKVFSIFPHKDCGDGVCISDLVLNQLPATQOQPIVSNQNK 810
QY 800 EFNVTVRNMGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESAST-EVSG 858
Db 811 RLTFVQLKNKKEASVNTETVVDVSENLF-----ASWMPVDGTEVTQOIAS 858
QY 859 ALKSTCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTQFOL 918
Db 859 SQKSVTCNGVTPALKSKQQTFTTFNFDNLQ-LNQOASISFALSESQENMADNSVL 917
QY 919 ELPKYAVYVMTVSHGVSTKYLNTASENTSRVMOHQYQVSNLQOR-----SLPI 968
Db 918 KLSLLYDAEIHIT-RSTNINFEVSLGNVSSV-HSFE--DIGPKFISIKVTGSGVPV 973
QY 969 SLVFLVRLNQTVWRDPQVTFSEN--LSSTCHTKE-----RLPSHS 1009
Db 974 SMA-----SVIIHIPQVTKDKNPLMYLTGVHTDQAGDISCEABINPLKIGQTSVV 1024
QY 1010 DELAE-LRKAPVNCSTAVCORIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNLLIV 1068
Db 1025 SFKSENPRHIELNCRATSCSNIMCWLDIQVGEYFLNVSTRIWNGTFAASTFTQVOLT 1084
QY 1069 STAEI-LFNDVSFTL-----LPQOGAFVRSQTETKVEPPE-VFNPLPIVGSVSGLL 1120
Db 1085 AABIDYNYQIYIEENTVITP-----LTINKPHEKVEVPTGVIVGSVIAGILL 1134
QY 1121 LALITAALYKLGFPKQYKDM 1141
Db 1135 LLALVAILWLKGLFPKRYKYM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Lar submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but no

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:9473098; PIDN:CAA82877.1; PID:9473099

C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F:169-344/Domain: von Willebrand factor type A repeat homology <WMA>

Query Match 18.3%; Score 1089; DB 2; Length 1178;

Best Local Similarity 28.1%; Pred. No. 6.5e-67;

Matches 348; Conservative 209; Mismatches 492; Indels 188; Gaps 44;

QY 5 VLLLTALTLCGHFNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIIVAAQR 60
Db 15 LMLVQGIILNCLAYNVGLFGAKIFSGPSEQFGYSVQQLTNPQGNLVLVGSMPSPENRM 74
QY 61 GSLYOC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQ 111
Db 75 GDVYKCPVDLPTATCEKUNLQNSASISNVTETKTNMSLGLTLTRNPGTGGFTCGPLWAH 134
QY 112 TCSNTYVKGCLFLFGSLNRQOPQ---KPEALRCQPOEDSDIAFLIDGSGIIPHDPFR 168
Db 135 QCGNQYATGIC---SDVSPDFQFLTSPFAVQACPSL-VDVVVVCDSESNIIY-WEA 187
QY 169 MKEFVSTVMEQLK--KSKTSLFMOYSEFRHFTFKFONNPNRSLVKPITQLG-RT 225
Db 188 VKNFLVKVFTGLDIGPKKTOVALIQYANEPRIIFNLNDFETKEDVMQATSETROHGGDLT 247
QY 226 HTATGIRKVVRELFNITNGARKNAFKILVITDGBKFGDPLGYEDVPEADREGVIRVI 285
Db 248 NTFRAIEFARDYAYQTSQGRPCATKVMVVVTDGESH-DGSKLKTIVIOCNDEILRFGI 306
QY 286 GV-----GDARFSEKSRQELNITIASPPRDHVFQVNNFEALKTIONQREKIFAIEGTQT 340
Db 307 AVLGYLNRNALDTKNLKEIKAIASPTERYEFNVADEAALLKAGTLGEOIFSIEGTQV 366
QY 341 GSSSFHEHMSQEGFSA--AITSNGPLSTVGSYDWAGVFLYTSKESSTFINMT--RVD 396
Db 367 G-GDNFQEMAQVGSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAPDQVL 425
QY 397 SDMN-DAYLYGAAAAILNRVQSLVGLAPRYQHIGLVAMFRONTQWGESNANV----KGT 451
Db 426 QDRNHSSFLUGYSVAALISTEDGVHVFAGAPRANYTGQVLYSVNK---QGNVTVIOSHRGD 482
QY 452 QICAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTR--GGQVSVCPPLPRQARWQCDVAV 509
Db 483 QIGSYFGSVLCSVDVDKDTITDVLVAGAPTMYNDLKEEGKVYLFITIKGILNQHQ---F 539
QY 510 LYGEQQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHS 569
Db 540 LSGPECTGNARFGSAIALSDINMDGFNDVIVGSPVENENSGAVIYNGHQT-IRTKYS 598
QY 570 QRIAGSKLSPR--LQYFGOSLGGQDITMDGLVLTGAQGHVLLRLQRSQVLRVKAIMEF 627
Db 599 QKITLNGAFRAHLOFFGRLSDGYDLNGDSITDVSIGALGQVQLWSSQSIADVAIEALF 658
QY 628 NPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVTDLALD----S 683
Db 659 TP-----DKITLLNKQAKITLKLCFRAEFPAQONNQV--AILFNMTLDADGHS 705
QY 684 GRPHSRAVFNETKNSTRTOVLGLTQTCET--LKLQLPNCIEDPVPSPVLRNLNFSLVGT 741

Db 706 SRVTSRGVFNSENRFQKNNVNVKQSCSHHSIQKPS---DVNPLDLRVDISLENP 762
 Qy 742 PLSAFGLNRPVLAEDAKLFTALPFPEKNCNDNICODDLSI-----TFSPMSLDC 792
 Db 763 GTS-----PALEYSETVKVSPFPFYKEGSDGICISDLILDVQQLPAQTQSF----- 811
 Qy 793 LVGGPREFNVTVTRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESAS 852
 Db 812 IVSNQKRLTFESVILKNRGESAYNTVLAEPSENLF-----ASFSPVVDGTE 859
 Qy 853 ST-EVSGALAKSTCSINHPFENSEVTFNITFDVDSKASIGNKLLKANVTSENWNPRT 911
 Db 860 VTCEVGSQSVTCVGVGPAKLSQQVTFNINFDNLQ-NLQNAANFQAFSSSQ--ET 916
 Qy 912 NKTE--FOLELPVAVVMVVTSHGVSTKYLNFTASNTSRVMQHQVQVSNLQNR----- 964
 Db 917 NKADNSVSLTIPLYDAELHLT-RSTNINFEIISDENAPSVIK---SVEDIGKPFISL 972
 Qy 965 -----SUPISLVFLV-----PVRLNQTVIWD 985
 Db 973 KVTAGSAPVSMALVTIHIPQVTEKNPLLYLTGTQDQAGDISCTAEINPLKLPHTA--- 1029
 Qy 986 RPQVTF-SENUSSTCHYKRLPSHSDFLAELKAPVNCSTAVCORIQCDIPFGIQEFP 1044
 Db 1030 -PSVSFKENFR---HTKE-----LDCRTTSCSNITCMLKDLHMAEY 1068
 Qy 1045 NATLKGNSLPDWYIKTSHNHLIYSTAEILFNSDVFLLPQGAFFVRSQTKETKPEPEVP 1104
 Db 1069 FINVTTRWNRTFAASTFQTVOLTAABEIDTHNQLFVIEENAVTIPLMIMKPTKAEVP 1128
 Qy 1105 NPLPLVIGSSVGGLLALALTAALYKLGFFKRYQKDM 1141
 Db 1129 T-GVIIGSIAGILLALLAMTAGLWKLGLFFKRYQKDM 1163

RESULT 10

A33998
 N: Integrin alpha-2 chain precursor - human
 C: Species: Homo sapiens (man)
 C: Date: 30-Mar-1990
 C: Accession: A33998; B56793; A53117
 J: Takada, Y.; Hemler, M.E.
 J. Cell Biol. 109, 397-407, 1989
 A: Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
 A: Reference number: A33998; MUID: 89308879; PMID: 2545729
 A: Accession: A33998
 A: Molecule type: mRNA
 A: Residues: 1-1181 <TAK>
 A: Cross-references: GB: X17033; NID: g33906; PIDN: CAA34894.1; PID: g33907
 A: Note: The authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
 R: Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
 Biochem. J. 279, 419-425, 1991
 A: Title: Separation of important new platelet glycoproteins (GP1a, GP1c, GP1c*, GP1Ia and
 A: Reference number: A56793; MUID: 92061944; PMID: 1953640
 A: Accession: B56793
 A: Molecule type: protein
 A: Residues: 30-43 <CAT>
 A: Experimental source: platelet
 R: Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
 J. Biol. Chem. 269, 463-469, 1994
 A: Title: The human alpha-2 integrin gene promoter. Identification of positive and negative
 A: Reference number: A53117; MUID: 94103255; PMID: 8276836
 A: Accession: A53117
 A: Molecule type: DNA
 A: Residues: 1-16, 'V', 18-21 <ZUT>
 A: Cross-references: GB: L24121; NID: g400342; PIDN: AAA16619.2; PID: g4583535
 A: Note: authors translated the codon GTA for residue 17 as Leu
 C: Genetics:
 A: Gene: GDB: ITGA2; CD49B
 A: Cross-references: GDB: 128031; OMIM: 192974
 A: Map position: 5q11.1-5q11.2

C: Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
 C: Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
 F: 1-29/Domain: signal sequence #status predicted <SIG>
 F: 30-1133/Domain: extracellular #status predicted <EXT>
 F: 172-347/Domain: von Willebrand factor type A repeat homology <WMA>
 F: 1134-1154/Domain: transmembrane #status predicted <TM>
 F: 1155-1181/Domain: intracellular #status predicted <CYT>
 F: 105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalen

Query Match 18.3%; Score 1087.5; DB 2; Length 1181;

Best Local Similarity 27.3%; Pred. No. 8.3e-67;

Matches 339; Conservative 214; Mismatches 498; Indels 193; Gaps 44;

Qy 3 LRVLTLTALT-----LCHGFNLDTENAMTFQ-ENARGFGQSVVQL---QSGRVVVGAPQE 53

Db 11 LPLLLVLALSQGILNCCLAYNGLPEAKIFSGSPSEQGYAVQOQFNPKNMILLVGPWS 70

Qy 54 IVAANGSILYQC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSLAATSPPOLLA 104

Db 71 GFPENRMGDYKCPVDLSTATCEKLNLTQSTSPNVTETKNTNMSLGLILTRNMGTTGGFLT 130

Qy 105 CGPTVHTCSENTVYVKGCLFGLFGLNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPH 164

Db 131 CGPLMAQCGNQYTTGVCSDISDPDF-QLSASFSPTQPCPSL-IDVVVCDSESNISYPW 188

Qy 165 DFRMKFEFVTVMEQLK--KSKTLFSLMYSEEFRIHFTPEKFNQNPNSRLVKPITQLL 222

Db 189 D--AVKNFLEKFGVGLDIPGTTKTOVGLIQYANNRVVFNLTNYTKKEMIVATISQTSYG 246

Qy 223 G-RTHATGTRKVVRELFNTNGARKNAFKILVITDGEKFGDPLGVEDVIEPADREGVI 281

Db 247 GDLTNTFGATQYARKYASAGRRSATKVMVWVTDGESH-DGSMUKLKAVIDQCNHNL 305

Qy 282 RYVIGV-----GDAFRSEKSRQELNTIASKEPRDHVFQVNNFEALKTIQNLREKIFAIE 336

Db 306 RFLAVLGVNLRNALDKNLIKETKATASIPTEFFNFNDEAALLKAGTLGQIFISIE 365

Qy 337 GTGTGSSSFEHMSQEGFSAAITSNGP--LLSTVSGYDMAGGVFLYTSKEKSTFINMT- 393

Db 366 GTVQG-GDNFQEMSVQFGFADYSSQNDILMLGAVGAFGWSGITVOKTSHGLIFPKQAF 424

Qy 394 -RVSDMN-DAYLGYAAAAIILNRVQSLVIGAPRYQHIGLVAMPFRONTGMWESNANV--- 448

Db 425 DQILQDRNHSYLGYSVAAISTGESTHFVACAPRANTGQIVLYSVN-----ENGNTIVI 479

Qy 449 ---KGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTR--GGOVSVCPPLRGQRAR 503

Db 480 QAHRGDDQIGSYFGSLCSVDVDKDTITDVLVVGAPMYMSDLKKEEGRVYLFTIKKGLGQ 539

Qy 504 WQDAVLVYGEQGPWGRFGAALTVLGDNVNGDKLTDVAIGAPGEDNDRGAVYLFHGTSGSG 563

Db 540 HQ---FLEGPEGIENTFGSAIAALSINMDGFNDVVGSPLENQNSGAVIYNGHQT- 595

Qy 564 ISPSHSQRIAGS--KLSPRIQYFGQSUGQDLMGDLVLTGVAQGHVLLRSQPVLRV 621

Db 596 IRTKYSQKILGSDGAFRSHLQYFGSLDGYGLNGDSITDVSIGAGQVQVQLWSQSADV 655

Qy 622 KATMEFNPREVARNVFECNDQVYVKGKEAGVRVCLHVQKSTRDLRGLRGQTSVVTYDLAL 681

Db 656 AIEASFTPEKI--TLVNKNAQII-----LKLCF-----SAKFRPTKONNOVAIVYNI 702

Qy 682 D----SCRPHSAVFNETKNSRROTQVGLGTQTC--ETLKLQPNCIEDPVSPIVLRNLN 735

Db 703 DADGFSRVTSRGLFKENNERCLOKNNVNAQSCPEHIYIQEPS---DVVNSLDLRVD 759

Qy 736 FSLVGTPLSAFGLNLRPVLAEDAQRLEFALPFFPEKNCNDNICODDLSITF-----SPMSLD 791

Db 760 ISLENPCTS-----PALEYSETAKVFSIPFHKDCGEDGLCISDLVDLRQIPAAEQBP 813

Qy 792 CLVVGGRPREFNVTVTRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESA 851

Db 814 FIVSNQNKRLTFSVTLKNKRESAYNTGIVVDFSENLF-----ASFSLPVDGT 861

Qy	852	SST-EVSGALKSTSCSINHIPIFPENBEVFTNITFDVDSKASLGNKULLKANVTSENMPR	910
Db	862	EVTCQVAASOKSVACDVGYPALRKQOVTFITNFDENLQ-NLONQASLFOALSESQEBN	920
Qy	911	TNKTEQLELPVKYAVYVUVTSHGCVSTKYLYNFTASENTRVMOHQYOVSNLGOR	964
Db	921	KADNLVNLKIPLLYDAEI---HLTRSTNINFYEISDGNVPSIVHSFEDVGPKFIISLK	976
Qy	965	---SLPISLVLFLV---PVRNLNQTVIWDR	986
Db	977	VTTGSPVPSMATVIIHPIQVTKERNPLMYLTGVOTDKAGDISCNADINPKIQOT	1032
Qy	987	PQVTF--SENLSSTCHTKERLPSSDFSFLAEURKAPVNVNCSTAVCQRIQCDDIPFGIOEEN	1045
Db	1033	SSVSFKSENR---HIKE-----LNCRTASCNSVTCMLKDVHMKGEYF	1072
Qy	1046	ATLKGNLSDFWYIKTSHNHLILVSTABI-LFNDVSFTLLPGQAFVRQSTETKVEFEVP	1104
Db	1073	VNVTTIRWNGTFASSTFQTVQLTAAAEINTYNEIYVI-----EDNTVTIPLMIM	1122
Qy	1105	NP-----LP--LIVGSSVGGELLALITALYKLGFPPKROYKDM	1141
Db	1123	KPEKAEVPTGVJIGSTIAGILLALLVALIWLKLGFFKRYEKM	1166

RESULT 12

A41131

A41131 lymphocyte-pegver's patch adhesion molecule alpha 4 precursor - mouse

lymphocyte-peyer's patch adhesion molecule

N; Alternate names: integrin alpha-4
C; Species: Mus musculus (house mouse)

C;Species: Mus musculus (house mouse)
C;Date: 05 Jun 1992 #sequence revision 05-Jun-1992 #text change 24-Sep-1999

C;Date: 05-Jun-1992 #sequence_revisio

C;Accession: A41131; S16742

R; Neuhaus, H.; Hu, M.C.T.; Hemler, J. Cell Biol. 115, 1149-1158, 1991

RESULT 11

A35854

A35634 integrin alpha-1 chain precursor - rat

C-Species: *Rattus norvegicus* (Norway rat)
 integrin alpha-1 chain precursor - rat

C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence revision 13-Sep-1991 #text change 20-Sep-1999

C; Date: 23-Oct-1990 #sequence
C; Accession: A35954. C11243

C;Accession: A35854; S11243
P. Ignatius M.J.: Large T.H.: Houde M.: Tawil J.W.: Barton. A.: Esch. F.: Carbonetto,

R; Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, I. Cell Biol 111 709-720 1996

J. Cell Biol. 111, 769-720, 1990

A;Title: Molecular cloning of the rat integrin alpha-1A
Accession number: M20954
MIM: 60328135
PMID: 2380249

A; Reference number: A 15054

A;Accession: A35854

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1180 <IGN>

A;Cross-references: GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query Match	Score 1082;	DB 2;	Length 1180;
18.2%			

Query Match
Best Local Similarity

BEST LOCAL SIMILARITY	27.7%	FREQ. NO:	28-00,
MATCHES	351:	MISMATCHES	484:
CONSERVATIVE	199:	INDELS	234:
GAPS		GAPS	47:

OV 6 I.I.I.TA.I.TI.CHGENI.DTENAMTEOENARG-EGOSVVOI.---OGSRVVVGAPOEIVAAANORG 61

6 LLLTALTLC HGFNDTENAMIFQENARG-FGQS VVQL---QGSRVVVGAFQVEI VAGNQK 31

DB 18 LTTVLGFCVSFNVDVKNMSFSGPPVEDMFGYTVQQYENEEGKNVLLIGSPVLVGQPNARIG 77

Q. 23. SYVOCDVSTCCCB-TPICUIRVEA-----IAMSICISIAATSTBPOLACGPTVHO 11

[illegible]

Db 78 DVYKCPVGRERAMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNGGFLACGPLYAY 13

Qy 112 TCSENTYVKGLCFLFGSNLRQQPQKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKE 17

[illegible]

.....

QY 172 FVSTVMEQLK--KSKTLFSLMQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLG-RHTTA 22

Db 192 FLNDLLKRMIDIGPKQTQVGIVQYGENVTHEFNLNKYSSTEEVLVAANKIGRQGGLOTMTA 25

1

QY 229 TGIRKVVRELFNITNGARKNAFKILWITDGEKFGDPLGYEDVIPEADREGVIRYVIGVG 28

Qy	59	----	QRGSLYQCDYSTG	----	SCEPIRLQVPVEAVNMISGLSLAATTSPPOLLAGCVTHQ	111	
			: : :		: : :		
Db	86	ASVINP	GAIVRCRIGKNPGQT	CQLQGLSP	-----NGEP-----CG----	K 122	
Qy	112	TCSENTV	VKGLCFLFGSNLRQOKPPEALRGCPQEDSDIAPLIDGSGSI	IHPDPRMKE	171		
			: : :		: : :		
Db	123	TCLEERD	NQWL----	GVTLSRQGE	-----NGSI-----	147	
Qy	172	FVSTVME	QKSKTFLFSLMOYSEPR	IHTFKEFQNNPNRSLVKPITQLLGRTH	TATGI 231		
Db	148	----	----	----	----	147	
Qy	232	RKVUREL	FNITNGAR-KNAPKIL	VITDGEKFGDPLGYEDV	IPBEADREGVIRVYVIGVDA 290		
			: : :		: : :		
Db	148	----	VTCGRMKN	IFYI	-----KNENKLEPTGCGY-----	174	
Qy	291	FRSEKSR	OELENTIASPPRDRH	FQV--NNPEALK	TIQNLREKIFAIEGTQTGSSSFEHE 349		
			: : :		: : :		
Db	175	--PPDL	TELSKRIAPCYQD	VVKFGENFAS	-----	203	
Qy	350	MSQEGF	SAATISNGPLLS	TVGSDWAGGVLY--TSKEKST	FINNTRVDSMDNDAYLG 406		
			: : :		: : :		
Db	204	-CQAGIS	FYTKOLLVNGAPGSSVY	TGSLFVYNNITNKYKAFDKQNVKF--	---GSLGY 259		
Qy	407	A--AAI	ILNRVQSLVLAGPRYOH	IGLVANFRQNTGHWESNAV--	---KGTQIGAYFGAS 460		
			: : :		: : :		
Db	260	SVGAGH	FRSQHTEVVGGAPE	QIGKAYIF----SIDEKEL	ILHEMKKKLGSYFGAS 315		
Qy	461	LCSVDV	DSNGSTDLVLGAPHY	EQTRGGQVSCPLPRGORARWQC--	DAVLYCEQOPWG 519		
			: : :		: : :		
Db	316	VCADV	LWADGFSU--LVGAP	MQSTIREEGRFVY--INSGCA	VNMAMETNLVGS DKYA-A 372		
Qy	520	RFGAAL	TVLGDVNGDKLTDVA	IGAPGEDNRGAVYLFHGTS	SGSISPHSQRIAGSKLSP 579		
			: : :		: : :		
Db	373	RFGESI	VNLGDIONDGEDVA	IGAPQEDDLQGA	IYVNGRA-DGISSTFSQRIEGLQISK 431		
Qy	580	RLQYFG	QSLSGGQDLTWG	LDLTVGA--OGHVLL	LRQSPVLVRKAIMENPREVARNVF 637		
			: : :		: : :		
Db	432	SLSMFG	QSIGQIDADNNGY	VADVAVGAFRSDSAVLL	TRPVIADVDSL--HPESVNR	TKF 490	
Qy	638	ECNDQV	VKGGEAGEVRVCLH	VOKSTRDRLREGQIQSV	VTYDLALDSGR-----PHSRAVF 692		
			: : :		: : :		
Db	491	DC-----	VENGWPS	VCIDLTLCFSYKGE	VPQYIVLFYNNSLDVNRKAESP	PRYFSS 543	
Qy	693	NETKN	STRROTQVLGLTQ	CETLKLQUPNCIEDP	SPSVILRNLFSVLGTPLS-----AFG 747		
			: : :		: : :		
Db	544	NGTSD	VITGSIQVSSREANC	RTHQAFMRKDVRLD	LTPIQIEAAHYLGHVIXRSTEEFP 603		
Qy	748	NLRVL	AEADQR-LFTAL	FPEKNCNDNI	QDDLSIT--FSRM-----SLDCLV	VGGPRE 800	
			: : :		: : :		
Db	604	PLQI	LOQKKEKIMKKTIN	IFARCAHEN--CSADL	QVSAKIGFLKPHENKTYL	AVGSMKT 662	
Qy	801	PNVT	TVRNDGEDSYRTQ	VTFPPDL	LSYRKVSTLONQORSWR	LACESASSTEVSGAL 860	
			: : :		: : :		
Db	663	LMUN	VSUFNAGDA	YETTLHVKUPV	GLYFIKILEEBK-----QINCE---	VTDSGVV 713	
Qy	861	KSTSCS	INHPHIFPENSE	VITNITPDVDSKAS	LGKLLKANVTSEN--NMPRN	TKTEFQ 918	
			: : :		: : :		
Db	714	Q-LDCS	IGIYIVDHLSD	ISIDISFLD	VSSLSRAEEDLSITVHATCENE	EDMDNLKHSRVTV 772	
Qy	919	ELPV	KYAVMVVTS	SHGVSTKYLNFT	ASENTSRV-----MOHQOV	QVSNLQORSIP-LSLVF 972	
			: : :		: : :		
Db	773	AIPLK	YEVKLTVHG	FVNPPTS	FVYGSNDENEPETCM	WEKMLTFHVINTGSMAP	NSVSEI 832
Qy	973	LVP	RLNQT	VIWDRPQV	TFSENL-----SSTCH	TKERLPSHSDFLAELRKA	PVNCVIA 1026
			: : :		: : :		
Db	833	MVPNS	FS-----	PQTDK	LFNILDVQTTGECH	FENYQ-----R 865	
Qy	1027	VCORI	QCDIPFFG	IOEBFNATL	KNLSFDWYIKTSHN	LLIVSTAE-----ILFN-----	1076
			: : :		: : :		
Db	866	YCALEQ	-----	QKSA	MTLKGIVR-----LSK	TRKLLYCYIKADPHCLN	FLCNFCGME 914
Qy	1077	----	DSVFTLLPGOGA	FVR	SQTETK	VPFEV-----PNP-----	1106

[illegible]

```
Db 648 ALGAVKQISLNISISNLGDDAYDANVSFNVSRELFF-----INMWQKEENGISCLELLES 701
Qy 854 TEVSGALKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNK 913
Db 702 DFL-----KCSVGFPMRSKSYEFVFDTSHLN-GEVEELSVFVTAQSG--NTER 750
Qy 914 TE-----FOLELPVKVAVMVVTS-----HGVSTKVLNFTASENTS---RVMQHQY 956
Db 751 SESLHDNTLVLMVPLMEHVDTSITGIMSPTSFVVGESVDAANFIQLDDLECHFQPINITL 810
Qy 957 QVSNLQORSPLISLVLFPVRLN-----QTVINDRPOVTFSENLSSTCHTKYER 1004
Db 811 QVNTGPTSTLPGSSVSISFNNRLSSGGAEMFHVQEMVVGQKNCSPQKNPTPCIIPOEQ 870
Qy 1005 LPSHSDFLAELRKA-----PVMNCSTAVCQRIQCDIPFFGQIEEFNATLGNLSFD 1055
Db 871 ENIHTTIFAFPTKSGRKVLDCFKGIGISCLTAHCN-----FSALAKEESRTI-----D 917
Qy 1056 WYIKTSHNHLIIVSTAELFNDSVFTLLPGQCAFVRQSTETKVEP-----1100
Db 918 IY-----MLLAT-EILKDDSSVQ-----FMSRAKVKVDPALRVVVEIAHGNPEEV 962
Qy 1101 ---FEVFN---PLPLIVG-----SSVGGLLLLALITAAALYKLGFFKQYKXDM 1142
Db 963 TVVFEALHNLPRGVVVGWIIAISLLVGLIFLLAVLLVLLWQGFRRRYKEII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.3%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 4.7e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 352 QEGFSAITNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----DNNDAYLG 405
Db 181 QAGFSGIIFSDNSALVMGAPGSYYLQGIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
Qy 406 YAAAI--ILNRVQSLVLGAPRYOHI-GLVAMFRONTGMWESNANVKGTOIGAYFGASLC 462
Db 240 YSLALGDFNGGVQDYVVGTPRAESLMGLVAIFDNLNQFN---QVMGTQIVAYFGYSVT 296
Qy 463 SVDVDSNGSDTLVLIGAPHYVEQTRGGQSVSCLPRGORARWQCDVLYGEQ-----514
Db 297 VVDI--NNDYDLDLVGAPMYMDGPAIQ-----RWEAGAVVYVLQNPDPVPGA 343
Qy 515 -----GQFWRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVYLFHGTSGS 562
Db 344 SNRLSSTLTGGQIRSRFGILSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG--SAN 402
Qy 563 GISFSHQRTAGSKLS-PRLOYFGQSLSGGODLTMDGLVLTVCAQ--GHVLLRSQPVL 619
Db 403 GLKSTPAQVLTPTLGHSGITTFGFSLQGGQMDKKNYPDILLVGAESANTAVLIRTPV 462
Qy 620 RVKAIMFNPREVARNVFECDNQVVVKGEAGEVRVCLHVQKSTRDLRREGIQSV-----674
```

```
Db 463 SLDATLNTSP-----IGINLENKTYE-LADGTMTVSTFIAMT 497
Qy 675 -----VTYDLALDSG-RPHSRAVENETKNSRRROTQVLGL-TQTCETLKQ 718
Db 498 CFYTYGNYLPDHDIDISYTVTVDSGIIANRRAMFVDMDSIITKRRLAVSTQCFDPLRAY 557
Qy 719 LNCIEDPVSPIVLRNLFNSLVGTPLSAFGN-----LRPVLAEADAQRLFTALPFFEK 769
Db 558 VGNISIEDKLTPIKVTIQLYDL-----NNDSESRLOPHEILPIIDNATMSTQTKVSION 609
Qy 770 NCNDNI CODDLSITFSFMSLDCLVGGRPREFNVTVVRNDGDSYRTQVTFFFFLDLSY 829
Db 610 NCVN-NICIPDLDTVT-PNLPNIVIGQTELTDVSLNRRGEDAFQSSLSVYPLQLQF 667
Qy 830 RKVSTLQONORSQSRWLACESASSTVSGALKSTCSINHPIPEN-----SEVTENIT 883
Db 668 VRU-----ERKANMDFSVTCSDD-----LRITCDTGNPMVGKNILEFGLTSLTFQVS 717
Qy 884 FDVDS-----KASLGNKLLKANVTSENMMPTNKTEFQLELPVKVAVMVVTSHGVS TK 938
Db 718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDTCLKLLSASYPEIVMYSTOED 775
Qy 939 YL-----NFTASENTSRVMQHOYQV-----SNLQORSLPIS-----LVFLVPV 976
Db 776 YVVPFPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQMPQKNEDGEYLFYLLGI 834
Qy 977 RLNQTVINDRPO-----VTFSENLSSTCHTKERLPSHSDFLAELRKA PVV 1021
Db 835 MTEEGVTCQLTQOKANPEGVKLEPSTKAKLSNSETTQVSGRKRREPEVAEALQTDN--VI 892
Qy 1022 NCSIAVCQRIQCDIPFFGQIEEFNAT-----LKGNLSFDWYIKTSHNHLIIVSTAELF 1075
Db 893 YCASDSCVLNCTI-----DEINASKSVKVRILGRF---W-----ERTF 928
Qy 1076 NDSVFTLLPGQGAFFRSQSTETKVE--PFEVNP-----LP-----1108
Db 929 QKAVSELTPVQVQATIASSASAAVKTIPIYNIPLPRDFSSTKASTLVTTTELVPVPTPIAW 988
Qy 1109 -LIVGSSVGLLLALLITAAALYKLGFFKR-----QYKDMMS--EGGPP 1148
Db 989 WIIIVSVLGGIILLIILGLWKCGFFERKKPGEKEEYAPVASADKGGPP 1038
```

Search completed: November 25, 2003, 14:21:50
Job time : 21.1684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.30328 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGFNLD.....FKRQYKDMSEGGPPGAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5940.5	99.7	1152	ITAM_HUMAN	P11215 homo sapien
2	4545	76.3	1153	ITAM_MOUSE	P05555 mus musculus
3	3503	58.8	1163	ITAX_HUMAN	P20702 homo sapien
4	3455	58.0	1162	ITAD_HUMAN	Q13349 homo sapien
5	1563	26.2	1170	ITAD_HUMAN	P20701 homo sapien
6	1249	26.0	1163	ITAM_MOUSE	P24063 mus musculus
7	1165.5	19.6	1167	ITAE_MOUSE	Q60677 mus musculus
8	1163	19.5	1179	ITAE_HUMAN	P38570 homo sapien
9	1108.5	18.6	1151	ITAI_HUMAN	P56199 homo sapien
10	1102.5	18.5	1189	ITAH_HUMAN	Q9UKX5 homo sapien
11	1099	18.5	1170	ITAE_BOVIN	P53710 bos taurus
12	1089	18.3	1178	ITAE_MOUSE	Q62469 mus musculus
13	1087.5	18.3	1181	ITAE_HUMAN	P17301 homo sapien
14	1085.5	18.2	1167	ITAI_HUMAN	O75578 homo sapien
15	1082	18.2	1180	ITAI_RAT	P18614 rattus norv
16	667	11.2	1039	ITAE_MOUSE	Q00651 mus musculus
17	638	10.7	1038	ITAE_HUMAN	P13612 homo sapien
18	630	10.6	1035	ITAE_HUMAN	Q13797 homo sapien
19	600	10.1	1032	ITAE_XENLA	Q91687 xenopus lae
20	571.5	9.6	1066	ITAE_CRISP	P17852 cricetidae
21	567.5	9.5	1053	ITAE_MOUSE	Q62470 mus musculus
22	555.5	9.3	1053	ITAE_MOUSE	P11688 mus musculus
23	550	9.2	1034	ITAE_CHICK	P26008 gallus gall
24	546.5	9.2	1050	ITAE_XENLA	Q08274 xenopus lae
25	540	9.1	1130	ITAE_HUMAN	P23229 homo sapien
26	535	9.0	1044	ITAE_MOUSE	P43406 mus musculus
27	532	8.9	1049	ITAE_HUMAN	P08648 homo sapien
28	531.5	8.9	1066	ITAE_HUMAN	P26006 homo sapien
29	530.5	8.9	1072	ITAE_CHICK	P26007 gallus gall
30	526	8.8	1048	ITAE_HUMAN	P06756 homo sapien
31	517	8.7	1044	ITAE_CHICK	P26009 gallus gall
32	512.5	8.6	1091	ITAE_MOUSE	Q61739 mus musculus
33	508	8.5	1396	ITAE_DROME	P12080 drosophila

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RT	J. Biol. Chem. 263:12403-12411(1988).			
RL				
RN	[2]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RL	[3]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RT	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RT	J. Immunol. 150:480-490(1993).			
RL	[5]			
RN	SEQUENCE OF 9-1153 FROM N.A.			
RP	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

ALIGNMENTS

34	498	8.4	1179	1	ITAE_MOUSE	Q61738 mus musculus
35	493	8.3	1033	1	ITAB_MOUSE	Q94000 mus musculus
36	491.5	8.3	1146	1	ITAI_DROME	Q24247 drosophila
37	489.5	8.2	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.2	1025	1	ITAB_HUMAN	P51708 homo sapien
39	486	8.2	126	1	ITAM_CAYPO	P11578 cavia porce
40	474.5	8.0	1226	1	PAT2_CABEL	P34446 caenorhabdi
41	472	7.9	1181	1	ITAE_HUMAN	Q13683 homo sapien
42	469.5	7.9	1106	1	ITAE_RAT	Q63258 rattus norv
43	454.5	7.6	1139	1	ITAI_CABEL	Q03600 caenorhabdi
44	424.5	7.1	1115	1	ITAE_DROME	O44386 drosophila
45	383	6.4	1000	1	ITAE_DROME	Q94000 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Paul H.L., Rosmarin A.G., Iken D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 78:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
RT conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Orvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3b. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; W76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RWHU1B.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

Query Match		99.7%	Score 5940.5;	DB 1;	Length 1152;
Best Local Similarity		99.9%	Pred. No. 0;		
Matches 1152; Conservative		0;	Mismatches	0;	Indels 1; Gaps 1;
QY	1	MALRVLLLTALTLCHGFNLDENAMTFQENARGFGSVVLOGSRVVVGAPOEIVAA	60		
DB	1	MALRVLLLTALTLCHGFNLDENAMTFQENARGFGSVVLOGSRVVVGAPOEIVAA	60		
QY	61	GSLYQCDXSTGCEPIRLQVPPVEAVNMSLGLSLAATTPPQLLAGCPTVHQTCS	120		
DB	61	GSLYQCDXSTGCEPIRLQVPPVEAVNMSLGLSLAATTPPQLLAGCPTVHQTCS	120		
QY	121	GLCFLFGSNLRQOPKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRKKEFV	180		
DB	121	GLCFLFGSNLRQOPKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRKKEFV	180		
QY	181	KSKTFLFSLMQYSEFRTHFTKEFONNPNPSLVKPIQLGRTHATGIRKVVREL	240		
DB	181	KSKTFLFSLMQYSEFRTHFTKEFONNPNPSLVKPIQLGRTHATGIRKVVREL	240		
QY	241	ITNGARKNAFKILVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAPFSEKSR	300		
DB	241	ITNGARKNAFKILVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAPFSEKSR	300		
QY	301	NTIASKPPRDHVFQNNFALKTIQNLREKIFAIEGTQTGSSSFHEHMSQEGFSA	360		
DB	301	NTIASKPPRDHVFQNNFALKTIQNLREKIFAIEGTQTGSSSFHEHMSQEGFSA	360		
QY	361	SNGPLLSTVGSVDMAGGVFLYTSKSTFINNTRVDSMDNDAYLGYAAIILRN	420		
DB	361	SNGPLLSTVGSVDMAGGVFLYTSKSTFINNTRVDSMDNDAYLGYAAIILRN	420		
QY	421	LCAPRYQHILGVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSD	480		
DB	421	LCAPRYQHILGVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSD	480		
QY	481	HYEOTRGQVSVCPPLRGORARWOCDAVLYGEOQOPWGRFGAALTVDVNGD	540		
DB	481	HYEOTRGQVSVCPPLRGORARWOCDAVLYGEOQOPWGRFGAALTVDVNGD	539		
QY	541	ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOYFGSLGGODL	600		
DB	540	ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOYFGSLGGODL	599		
QY	601	DLTVAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVR	660		
DB	600	DLTVAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVR	659		
QY	661	STRDLREGQIQSVVTDIALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCET	720		
DB	660	STRDLREGQIQSVVTDIALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCET	719		
QY	721	NCIEDPVSPIVLRNLSVGTPLSAFQNLRPVLAEDAQLFTALPFKNCNDNICQ	780		
DB	720	NCIEDPVSPIVLRNLSVGTPLSAFQNLRPVLAEDAQLFTALPFKNCNDNICQ	779		
QY	781	LSITFSMSLDCVAVGGPREFNVTVNRNDGSDSVRTQVTFEPDLDSYRKVST	840		
DB	780	LSITFSMSLDCVAVGGPREFNVTVNRNDGSDSVRTQVTFEPDLDSYRKVST	839		
QY	841	QRSWLACSSASTVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNK	900		
DB	840	QRSWLACSSASTVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNK	899		
QY	901	NVTSENMPRTNKTETQELPQVAVYVVTSHGVSTKYLNFTASNTSRVNHQVQ	960		
DB	900	NVTSENMPRTNKTETQELPQVAVYVVTSHGVSTKYLNFTASNTSRVNHQVQ	959		
QY	961	LGQSLPISLVLVPLVRLNQTWDRPQVTFSENLSSTCHTKERLPSPSHDFL	1020		
DB	960	LGQSLPISLVLVPLVRLNQTWDRPQVTFSENLSSTCHTKERLPSPSHDFL	1019		

QY	1021	VNCIAVCQRIQCDIPPFQIOEENATLKGNSLSDWYIKTSHNHLIIYSTAIEL	1080		
DB	1020	VNCIAVCQRIQCDIPPFQIOEENATLKGNSLSDWYIKTSHNHLIIYSTAIEL	1079		
QY	1081	TLLPQGAQFVSQSTETKVEPEVNPPLPLIYVSSVGGLLLLALITAAALYK	1140		
DB	1080	TLLPQGAQFVSQSTETKVEPEVNPPLPLIYVSSVGGLLLLALITAAALYK	1139		
QY	1141	MMSEGGPPEGAPQ 1153			
DB	1140	MMSEGGPPEGAPQ 1152			
RESULT 2					
ITAM MOUSE					
ID	ITAM MOUSE	STANDARD:	PRT:	1153 AA.	
AC	P05555;				
DT	01-NOV-1988 (Rel. 09, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MOI).				
GN	ITGAM.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89312584; PubMed=3044779;				
RA	Pyteia R.;				
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";				
RL	EMBO J. 7:1371-1378 (1988).				
RN	[2]				
RP	SEQUENCE OF 11-45 FROM N.A.				
RC	STRAIN=BA1B/c; TISSUE=Spleen;				
RX	MEDLINE=86287312; PubMed=2942940;				
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,				
RA	Larson R.S., Roberts T.M., Springer T.A.;				
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).				
RN	[3]				
RP	SEQUENCE OF 17-28.				
RX	MEDLINE=85188276; PubMed=3887182;				
RA	Springer T.A., Teplow D.B., Dreyer W.J.;				
RT	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";				
RL	Nature 314:540-542 (1985).				
CC	-!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO AN IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.				
CC	-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.				
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.				
CC	-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.				
CC	-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.				

CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC
 CC EMBL: X07640; CAA30479.1; --
 CC EMBL: M14293; AAA39484.1; --
 CC HSP; S00551; S00551.
 CC HSP; P11215; IABX.
 CC MGD; MGI:96607; Itgam.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 5.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Calcium; Repeat.
 FT SIGNAL 1 16
 FT CHAIN 17 1153
 FT DOMAIN 17 1105 INTEGRIN ALPHA-M
 FT TRANSMEM 1106 1129 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1130 1153 POTENTIAL.
 FT REPEAT 1 84 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 2 84 FG-GAP 1.
 FT DOMAIN 164 350 VWFA.
 FT REPEAT 337 400 FG-GAP 3.
 FT REPEAT 401 452 FG-GAP 4.
 FT REPEAT 454 515 FG-GAP 5.
 FT REPEAT 517 575 FG-GAP 6.
 FT REPEAT 580 632 FG-GAP 7.
 FT CA_BIND 465 473 POTENTIAL.
 FT CA_BIND 529 537 POTENTIAL.
 FT CA_BIND 592 600 POTENTIAL.
 FT SITE 1132 1136 GFFKR MOTIF.
 FT DISULFID 66 73 BY SIMILARITY.
 FT DISULFID 105 123 BY SIMILARITY.
 FT DISULFID 654 711 BY SIMILARITY.
 FT DISULFID 770 776 BY SIMILARITY.
 FT DISULFID 999 1023 BY SIMILARITY.
 FT DISULFID 1028 1033 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 772 772 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 801 801 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 881 881 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 941 941 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 994 994 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;
 Query Match 76.3%; Score 4545; DB 1; Length 1153;
 Best Local Similarity 74.5%; Pred. No. 6e-300;
 Matches 860; Conservative 141; Mismatches 151; Indels 2; Gaps 2;

QY 1 MALRVLLLTALTLCGHNLDTENAMTFOENARGFGQSVVQLGSGRVVYVGAPOEIVAAQNR 60
 Db 1 MTKALLVTALALCHGFLNLDTEHPMTFOENAKGFQNVVQLGTSVVVAAPOEAKAVQT 60
 QY 61 GSYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVK 120
 Db 61 GSYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVK 120
 QY 121 GLCFLFGSNLRQOPKPFPEALRGCCQEDSDIAFLDGSIIIPHFRMKERFVSTVMEQL 180
 Db 121 GLCYLFGSNLRQOPKPFPEALRGCCQEDSDIAFLDGSIIIPHFRMKERFVSTVMEQL 180
 QY 181 KSKTFLSLMOYSEBFRIHFTPEKFNPNPRLSVKPIQTOLLGRTHHTATGIRKVVRELFN 240
 Db 181 KSKTFLSLMOYSEBFRIHFTPEKFNPNPRLSVKPIQTOLLGRTHHTATGIRKVVRELFN 240
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPIEADREGVIRYVIGVDGDAFSEKSRQL 300
 Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPIEADREGVIRYVIGVDGDAFSEKSRQL 300
 QY 301 NTIASKPPRDHVQVNNFEALKTIONLBKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
 Db 301 NTIASKPPRDHVQVNNFEALKTIONLBKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
 QY 361 SNGPLLTSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
 Db 361 SNGPLLTSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
 QY 421 LGAPRYQHIHGLVAMFRQNTGMNESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
 Db 421 LGAPRYQHIHGLVAMFRQNTGMNESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
 QY 481 HYYETRGQGVSVCPPLRG-RARWQCEALLHGDQCHPWGRFGAALTVLGDVNGDKLTDVA 540
 Db 481 HYYETRGQGVSVCPPLRG-RARWQCEALLHGDQCHPWGRFGAALTVLGDVNGDKLTDVA 540
 QY 541 ICAPGEEDNRGAVLPHGTSGSGISPSHSORISAGSLSPRLQYFGQSLSGGODLTMDGLV 600
 Db 541 ICAPGEEDNRGAVLPHGTSGSGISPSHSORISAGSLSPRLQYFGQSLSGGODLTMDGLV 600
 QY 540 ICAPGEQNGQAVIIFYGASIASLSASHSHRIIGHAFSPQLQYFGQSLSGGODLTMDGLM 599
 Db 540 ICAPGEQNGQAVIIFYGASIASLSASHSHRIIGHAFSPQLQYFGQSLSGGODLTMDGLM 599
 QY 601 DLTVGAQGHVLLRLRSQPLVRVKAIMEFNPVERNARVNECNQDVVQKEAGEVVRVCLHVOK 660
 Db 601 DLTVGAQGHVLLRLRSQPLVRVKAIMEFNPVERNARVNECNQDVVQKEAGEVVRVCLHVOK 660
 QY 600 DLAVGAQGHVLLRLRSQPLVRVKAIMEFNPVERNARVNECNQDVVQKEAGEVVRVCLVRVK 659
 Db 600 DLAVGAQGHVLLRLRSQPLVRVKAIMEFNPVERNARVNECNQDVVQKEAGEVVRVCLVRVK 659
 QY 661 STRDRUREQIQSVVTVYDIALDSGRPHSRAVNETKXSTRQTQVGLTQTCTETLKLPL 720
 Db 661 STRDRUREQIQSVVTVYDIALDSGRPHSRAVNETKXSTRQTQVGLTQTCTETLKLPL 720
 QY 660 NTKDRLREGDIQSTVTYDIALDPVRSRIIRAFDETKNTRRRRTQVFLGMLQKCTELKLILP 719
 Db 660 NTKDRLREGDIQSTVTYDIALDPVRSRIIRAFDETKNTRRRRTQVFLGMLQKCTELKLILP 719
 QY 721 NCIEDPVSPVLRLNFSVLGTPLSAFGNLRPVLAEADAQRLFTALPFPEKNCNDNICODD 780
 Db 721 NCIEDPVSPVLRLNFSVLGTPLSAFGNLRPVLAEADAQRLFTALPFPEKNCNDNICODD 780
 QY 720 DCVDDSVSPILRLNLTVLGEPURSGNLRPVLAMDAQRAFFTAFFPEKNCNDNICODD 779
 Db 720 DCVDDSVSPILRLNLTVLGEPURSGNLRPVLAMDAQRAFFTAFFPEKNCNDNICODD 779
 QY 781 LSITFSMSLDCLVVGGPREFNVTVTRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
 Db 781 LSITFSMSLDCLVVGGPREFNVTVTRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
 QY 780 LSITFSMSLDCLVVGGPREFNVTVTRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRS 839
 Db 780 LSITFSMSLDCLVVGGPREFNVTVTRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRS 839
 QY 841 QRSWRL-ACESASTSEVSGALKSTSCINHPFPENSEVFNITFPDVSQKASLGNKLLK 899
 Db 841 QRSWRL-ACESASTSEVSGALKSTSCINHPFPENSEVFNITFPDVSQKASLGNKLLK 899
 QY 840 KKEWFKPAESSSSSEGHGALKSTTNINHPFPANSEVFNITFPDVSQKASLGNKLLK 899
 Db 840 KKEWFKPAESSSSSEGHGALKSTTNINHPFPANSEVFNITFPDVSQKASLGNKLLK 899
 QY 900 ANYTSENMPRTNKTFFOLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMOHOVQS 959
 Db 900 ANYTSENMPRTNKTFFOLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMOHOVQS 959
 QY 900 AIVASENNMSTRHKTFRQLELPVKYAYIWMIVTSESSIRYLNFTASENTRVMOHOVQS 959
 Db 900 AIVASENNMSTRHKTFRQLELPVKYAYIWMIVTSESSIRYLNFTASENTRVMOHOVQS 959
 QY 960 NLQORSLSPLVLPVRLNQTIVMDRPQVTFSENLSSTCHTKERLPHSDFLAELRKAP 1019
 Db 960 NLQORSLSPLVLPVRLNQTIVMDRPQVTFSENLSSTCHTKERLPHSDFLAELRKAP 1019
 QY 960 NLQORSLSPLVLPVRLNQTIVMDRPQVTFSENLSSTCHTKERLPHSDFLAELRKAP 1019
 Db 960 NLQORSLSPLVLPVRLNQTIVMDRPQVTFSENLSSTCHTKERLPHSDFLAELRKAP 1019
 QY 1020 VVNCSTAVCORIQCDIPFGIQEENFATLKNLSFDWYIKTSHNHLIVSTAILFNDSV 1079
 Db 1020 VVNCSTAVCORIQCDIPFGIQEENFATLKNLSFDWYIKTSHNHLIVSTAILFNDSV 1079
 QY 1020 VLNCSTAVCKRIQCCLPSTFNTQIFNVTLKGNLSFDWYIKTSHNHLIVSTAILFNDSV 1079
 Db 1020 VLNCSTAVCKRIQCCLPSTFNTQIFNVTLKGNLSFDWYIKTSHNHLIVSTAILFNDSV 1079

Qy 1080 FTLPGQAFVRSQTKVPEFVNPVLPIVSGSVGLLLALITAAKYKLGFFKQYK 1139
Db 1080 FALLPGQESVRSQTKVPEVHNVPVLIVSGSVGLLLALITAGLYKLGFFKQYK 1139
Qy 1140 DMNSEGPPGAEPO 1153
Db 1140 DMNNEAAPQADPO 1153
RESULT 3
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95.";
RL EMOB J. 6:4023-4028(1987).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN (3)
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN (4)
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFMA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR ENBL; M81695; AA59180.1; --
DR ENBL; Y00093; CAAG8283.1; --
DR ENBL; M29165; -; NOT ANNOTATED CDS.
DR ENBL; M29487; AAAS1620.1; ALT SEQ.
DR ENBL; M29482; AAAS1620.1; JOINED.
DR ENBL; M29483; AAAS1620.1; JOINED.
DR ENBL; M29484; AAAS1620.1; JOINED.
DR ENBL; M29485; AAAS1620.1; JOINED.
DR ENBL; M29486; AAAS1620.1; JOINED.
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; --
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163
FT DOMAIN 20 1107
FT TRANSMEM 1108 1128
FT DOMAIN 1129 1163
FT REPEAT 34 87
FT REPEAT ? ?
FT DOMAIN 165 351
FT REPEAT ? ?
FT REPEAT 402 453
FT REPEAT 455 517
FT REPEAT 518 576
FT REPEAT 581 633
FT CA_BIND 466 474
FT CA_BIND 530 538
FT CA_BIND 593 601
FT SITE 1131 1135
FT DISULFID 69 76
FT DISULFID 108 126
FT DISULFID 655 712
FT DISULFID 771 777
FT DISULFID 848 863
FT DISULFID 998 1022
FT DISULFID 1027 1032
FT CARBOHYD 61 61
FT CARBOHYD 89 89
FT CARBOHYD 392 392
FT CARBOHYD 697 697
FT CARBOHYD 735 735
FT CARBOHYD 899 899
FT CARBOHYD 939 939
FT CARBOHYD 1050 1050
FT CONFLICT 490 490
FT CONFLICT 756 756
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;
Query Match 58.8%; Score 3503; DB 1; Length 1163;
Best Local Similarity 61.3%; Pred. No. 2.7e-229;
Matches 700; Conservative 138; Mismatches 237; Indels 6; Gaps 4;
Qy 5 VLLTALTLCGHFNLDTENAMTFQENAGFGQSVVLQGGSRVVVGAPOEIVAAHQSGSLY 64
Db 8 LLLFTALATSLGPNLDTELTAPRVDSAGFGDSVVQVANSVVVVGAPQKITAANQTGGLY 67

Qy	65	QCDYSTGSCBPIRLQVPVPEAVNNSLGLSLAATTSPPQLLAGCPTVHOTCSENTYVKGLCF	124
Db	68	QCQYSTGACBPIGLQVPPEAVNNSLGLSLASTTSPQLLAGCPTVHHECRNNYLTGLCF	127
Qy	125	LFGSNLRQQKQKPEALRGCPQEDSDIAPLIDGSGSIIIPHDRRMKEFVSTVMEQLKSKS	184
Db	128	LLGPT--QLTORLPVSRQCPROQDILVFLIDGSGSISRNFATMNFVRAVISQORPS	185
Qy	185	TLFSLMOYSBEFRIHFTFKBFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFNWTNG	244
Db	186	TQFSLMQFSNKFOTHTFBEFRATSNPLSLLASVHQLQGTYTATAIQNVVHRLPHASVG	245
Qy	245	ARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVGDAPFREKSKRQELNTIA	304
Db	246	ARRDATKILVITDGKKEGSDLYKDVIPWADAAGIIRYAIGVGLAFQFNRNNSWKELENDIA	305
Qy	305	SKPPRDHVFQVNNFEALKTIONLREKIFAIEGTQGTSSSSFEHENSQEGFSAAITSNGP	364
Db	306	SKPSQEHIFKVEDFDALKDIONLQKKEKIFAIEGTETSSSFLEMAQEGFSAVTFDGP	365
Qy	365	LLSTVGSYDWAGGVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAP	424
Db	366	VLGAVGSFTWSGAFLYPNNKSTPFINMSQENVDMRDSYLGYSTELALMKGVQSLVLGAP	425
Qy	425	RYOHIGLVANFRQNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYYE	484
Db	426	RYQHTGKAVITQVSRQWRMAEYVGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYYE	485
Qy	485	QTRGGQVSVCPPLRGQBARQCDVAVLYGEOQPGWRFGAALTVLGVDVNGDKLTDVAIGAP	544
Db	486	QTRGGQVSVCPPLRGWR--RWCDAVLYGEOGHPWRFGAALTVLGVDVNGDKLTDVIGAP	544
Qy	545	GEEDNRGAVLPHGTSGSGTSPSHSQRIAGKLSPLRIQYFGQSLSGQDLTMQGLVDLTV	604
Db	545	GEENRGAVLPHFGVLGSPISPHSQRIAGSQLSSRLQYFGQALSGQDLTMQGLVDLAV	604
Qy	605	GAGCHVLLRSQVPLRVKATMENPREVANRVNCEINDOVYKGEAGEVRVCLHVQKSTRD	664
Db	605	GARQVULLKTRPVLWVGVSQMFIPAEIPSAFECREQVQVSEQTLVQSNICLYIDKRSKN	664
Qy	665	RLREGQIQSVVYTDLALDSGRPHSRVAFNETKNSRRTQTVLGLTCTCTKLQLPNCIE	724
Db	665	LLGSRDLQSSVTLDLADLPGRLSPRATFQETKNSRSRVRLGLKAHCENFNLLPSCVE	724
Qy	725	DPVSPVIRLNFSLVGTPLSAFGNLRPLVLAEDAQRLTALPPEKNGCNDNICODDLISIT	784
Db	725	DSVTPITLRLNFTLVGRPLLAFLAFLRLRMLAALAQRYPYASLPPEKNGCADHICQDNLGIS	784
Qy	785	FSFMSLOCLVVGGPREFNVTVTVRNDCEDSVRTQVTFEFPFLDLSYRKVSTLQNRORSRM	844
Db	785	FSFPLKSLLVGSNLENAEVMWVWDCEDSYGTTITFSPHAGLSYRVVASEGQKQQLRSL	844
Qy	845	RLACESASSTEVSGALKSTSCSINHPIFPENSEVFTNITFDVDSKASLGNKLLKANVTS	904
Db	845	HLTCDSAPVG--SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVIGDRLLLTANVSS	902
Qy	905	ENNMPRNKTKEFQLELPLVKYAVYVWVTSHGVSHTKYLNFATS-ENTSRVMOHQVQVSNLQG	963
Db	903	ENNTPTSKTTFQLELPLVKYAVYVWVSHEQFTKYLNFSESEKESHVAMHRYQVNNLQG	962
Qy	964	RLSPISLVPLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNC	1023
Db	963	RDLPSVINFWPVLNQBAVMMDVEVSHQNPQPSLRCSSEKIAPPASDFLAHIQKNPVJDC	1022
Qy	1024	SIAYCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTSHNHLITVSTAELFNDSVFTLL	1083
Db	1023	SIACLRFCRCDVPFSFQEELDFTLKNLSFGWVRQILQKKVSVSVAEITFDTSVYSQL	1082
Qy	1084	PGQAGFVRSQTETKVEFPEVNPPLPLVGSVGGLLLLALITALYKLGFFKQYKDMMS	1143
Db	1083	PGQAFVRAQOTTVLEKYKVNHPPLPLVGSIGLLLLALITALYKVGFFKQYKEMME	1142


```

Db      342 SFNMLSSSGISADLSKSHAVGVAGAKDWAGGFLDLREDLQGTVCQBELTSDVRGGY 401
Qy      404 LGYAAA-IILNRVQSLVGLGAPRYQHIGLVAMFR--QNTGMWESNANVKGTOIGAYFGAS 460
Db      402 LGYTVAMVTSRSSRLLAAGAPRYOHVQVLLFOAPEAGGRWNTQKTEGTIGSYFGGE 461
Qy      461 LCSVDVDSNGSTDLVLIIGAPHYEYTRGGQVSVCLPRGRARWOCDAVLVGEQOPWGR 520
Db      462 LCSVDLDQDGEALLIGAPLFFGEGRGRVFTY---QRROSLFEMVSELQDGPYPLGR 518
Qy      521 FGAALTVLGDVNGDKLTVAIGAPCEEDNRGAVLFGHTSGSISPSHSQRIAGSKLSPR 580
Db      519 FGAALTALTDINGDLRTDVAVGAPLEE--QGAVYIFNCKPG-GLSPQSPQRIQGAQVFP 575
Qy      581 LOYFQCSLGGQDLTMDGLVDLTGVAQGHVLLRSQPLRVKAINMEFNPVARNVFCN 640
Db      576 IRWFGRSHGVKDLGDLADVVVCAEGRVVVLSRPVVDVVTLSFSPPEEIPVHEVCS 635
Qy      641 DQVVGKENG-BVRVCLHVQKSTRDLREGQIQSVVTDALDSGRPHSRVAFNETKNT 699
Db      636 YGAREEQKHGVKLCACFRIKPLTPQ--FQGRLLANLSYTLQDGHMRSGLFPDGSHEL 693
Qy      700 RQQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNFSLV---GTPLSAFGN-LRPVLA 755
Db      694 SNTSITP-DKSCDLFHFFPICQDLISPINVSNLSEEGTPROCKGRAMQPILRP 752
Qy      756 DAQRLFTALFPPEKNCNDNICQDLSITFSFMSLDCLVGGP-----GPLRMSASLAVEWTL 804
Db      753 SIHTV-TKEIPPEKNCGEDKCEANLTLSPPARS-----GPLRMSASLAVEWTL 804
Qy      810 DGEDSYRTQVTFEFLDLSYRKVSTLQNRQSRWRLACESASTEVSGAL-KSTSCSIN 868
Db      805 SGEDAYVWRDLDFPRGSLFRKVEMLQ---PSHRMPVSCCEL--TEGSSLLTKLKCNVS 859
Qy      869 HPFIFENSEVTNITFDVDSKASLGNKLLKANVTSEN-NMPTNKTEFOLELPVYKAVY 927
Db      860 SPIFKAGQSVQVFMFTLLNSWEDFVELNQTVCHEENSSLSQEDNSAATHIPVLYPN 919
Qy      928 MVVTSHTGVTYKLTNTASENTSRVMOHQVQVNLGQSLPISLVLPVRLNQTIVMDRP 987
Db      920 ILTKQENSTLYISTPKGPKTQOVHVYQV-----RTPSAYDHNMPT-LEALVGVVRP 973
Qy      988 Q-----VTSENLSS-----TCHTKE-RLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFF 1038
Db      974 HSEDILITYTWSQTDPLVTCHESDLKRPSE---AEQPCLPGV-----QFRCPVIF- 1021
Qy      1039 GLOEFENATLKNLSFDMVTKTSHNHLIVSTAELFNDISVFTLLPGQAFVRSQETTKV 1098
Db      1022 --RWEILIQVGTVELSKIEIKAS-STLSLCSLSVSFNSSKHFHLYGSKA-SEAQVLVKV 1077
Qy      1099 EPFEVNPPLVIVGSSVGLLLALITAALYKLGFFKQYKQDM-SEGPPGCAEP 1152
Db      1078 DLIHEKEMLVVYVLSGIGGLVLLFLFLALYKVGFKRNLKEMKMEADGVNGSP 1132

RESULT 7
ITAE_MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7882170;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,

```

```

RA Kilshaw P.J., Weis J.H.;
RT "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VMFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U12236; AAC52142.1; -.
CC HSP: P11215; IABX.
CC MGD: MGI:1298377; Itgae.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWF_A.
CC Pfam: PF01839; FG-GAP; 3.
CC Pfam: PF00357; integrin_A; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS: PR01185; INTEGRINA.
CC PRINTS: PR00453; VWFADOMAIN.
CC SMART: SM00191; Int_alpha; 3.
CC SMART: SM00327; VWA_1.
CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE: PS02334; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Magnesium;
CC Calcium.
CC SIGNAL 1 19
CC CHAIN 20 1167
CC CHAIN 20 181
CC CHAIN 183 1167
CC DOMAIN 20 1114
CC TRANSMEM 1115 1137
CC DOMAIN 1138 1167
CC REPEAT ? ?
CC REPEAT ? ?
CC DOMAIN 149 192
CC DOMAIN 193 384
CC REPEAT ? ?
CC REPEAT 449 501
CC REPEAT 503 564
CC REPEAT 566 631
CC REPEAT 634 686
CC CA_BIND 514 522
CC CA_BIND 578 586
CC CA_BIND 646 654
CC DOMAIN 185 191
CC SITE 1140 1144
CC DISULFID 72 83
CC DISULFID 130 164
CC DISULFID 698 754
CC DISULFID 814 820
CC DISULFID 884 898
CC DISULFID 998 1023
CC -----
CC BY SIMILARITY.
CC INTEGRIN ALPHA-E.
CC INTEGRIN ALPHA-E LIGHT CHAIN.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC FG-GAP 1.
CC FG-GAP 2.
CC X-DOMAIN (EXTRA DOMAIN).
CC VWFA.
CC FG-GAP 3.
CC FG-GAP 4.
CC FG-GAP 5.
CC FG-GAP 6.
CC FG-GAP 7.
CC POTENTIAL.
CC POTENTIAL.
CC GLU-RICH (ACIDIC).
CC GFFKR MOTIF.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.

```

FT DISULFID 1031 1047 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 718 718 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 829 829 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 846 846 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 925 925 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFDD CRC64;
Query Match 19.6%; Score 1165.5; DB 1; Length 1167;
Best Local Similarity 29.1%; Pred. No. 8.6e-71;
Matches 362; Conservative 213; Mismatches 465; Indels 205; Gaps 43;
QY 5 VLLTALTLCGHNLDTEA--MTFQENARGFGOSVLOQSRVVVGAPQEIIVAAQNGS 62
DB 8 LNCWASLKPGAFNMDVDWAWTALQCAPAVLSLHLDPN-----NOTCLLVARRSS 62
QY 63 -----LYQCDYSTGSCPEIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHQT 113
DB 63 NRNTAALYRCAISI-SPEIACQ-PVEHICMPKGRYQGVTLVGNHNGVLCVQVQARKPR 120
QY 114 SENTYVGLCLFSGNLROQPKPEALRG-----C----- 144
DB 121 SLNSELTCACSLTPNLDLQQAQYFSDLEGFLDPCAHDVSDGYCRSKGSGTGBEKSARR 180
QY 145 -----POEDSDIAFLIDSGSIIPHDFRRMKEFVSTVMEQL--KSKTFLPSLMOYS 193
DB 181 RTVEEEDDEEDGETEIALVLDGSGSIGPSDFQKAKNISTMRNFYEKCFECFNALVQYG 240
QY 194 EEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAPKIL 253
DB 241 AVIQTEPDQSRDINASLAKVQSIQVQKVTKTASAMQHVLDFIPSGRSRKKALKVM 300
QY 254 VITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSGQELNTTASKEPRDHVF 313
DB 301 VVLTDGDIQDPLNLTTVINSPPKQGVVFAIGVGRFKNNNTYRELKLIASDPKEARTF 360
QY 314 QVNNFEALKTIONLRKIFALEGTQGTGSSSEFHEMSQEGFSAITSNRP-LLSTVGSY 372
DB 361 KVTNYSALDGLSLKIQRIQVHMEGT---VGDALQYLAQTGFSQIILDKGVLLGTVGAF 417
QY 373 DWAGGVFLY-TSKEKSTFINMT-RVDS-DMNDAYLGYAAAILNRNVQSLVGLGAPRYOHI 429
DB 418 NWSGALLYSTQNGRCFLNQTAKEDSRTVQSYLVGLSLAVLHKAHGISYVAGAPRHKL 477
QY 430 GLVAMFRONTGMWESNA---NVKGTQIAYFGASLCSVDVDSNGSTDLVLICAPHYEQT 486
DB 478 GAVFELRKEDR--BEDAFVRRIEGEQMGSGYFSGVLCVPDIDMDGTDFLLVAAPFYHIG 535
QY 487 RGGGVSVCLPRGQPARQCDVAVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAP-- 544
DB 536 BEGRVYVQVPE-QDASFSLAHTLSGHPGLTNSRFGFAMAAGVDINQDKFTDVAIGAPLE 594
QY 545 ---GEEDNRGAVLPHGTSGSIGSPHSQRIAGSKLSPRLQYFQSGLSGGQDLTMGLV 600
DB 595 GFGAGDGSYGVYIYNGHSG-GLYDPSQIRASSVASGLHYFGMSVSGLDLFGDGLA 653
QY 601 DLTGVAQGHVLLLRSPQVLRVKAINEFNPREVARNVFCNDQVVKGEAGEVRYCLHVQK 660
DB 654 DITVGSRDSAVVLRSPVVDLTVSMTFP-----DALPMVFITGKM--DVNLCFEVD 703
QY 661 S---TRDLRREGIQSVTYDLDALDSGRPHSRAVENETKNSRRQTQVGLGTQTC----- 712

DB 704 SVVASEPGLUREMFLNFTVDV-----TKQRQLQCESSGSCQLRKWN 748
QY 713 -----ETLKLQLPNCIEDPVPSPVLVLRNLFSLVGTPLSAGNLR-----PVLAED 756
DB 749 GSGFLCEHFWLISTBEL-----CEEDCFSNITIKVTE-----FQTSGGRRDYPNPTL--D 797
QY 757 AORLFTALF--PFERKNCNDNQCDDLSITFSFMSLDCLVVGPGPREFNVTVVRNDGEDS 814
DB 798 HYKEPSAIFOLPEYKCKKVFCAIEIQLTTN-ISOQLVVGVTKEVTMNSILTSNGEDS 856
QY 815 YRTQVTFPPDLVSRYKVTSTLQNRQSRWRACSSASSTEVSGALKSTSCSINHIPIPE 874
DB 857 YMTNMAINPRNLQFKKI-----QKPVSPDVCQDDPKPV---ASVLVNMCKIGHPI-L-K 906
QY 875 NSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTETFOLELPVKYAVYVMTVSHG 934
DB 907 RSVNVSVTWQLEESVFPNRTADITVTISNSNEKSLARETR---SLQFRAHIAVLNR-- 961
QY 935 VSTKYLNTASNTSRVMQHQYQVSNLQGRSLPISLVFLVLPVLRNQTIVWRPQVTFSEN 994
DB 962 PSVMYWN--TSQSPDHKEFFFNHVGENLFGAVFOQLQICVPKLODF-----QIVRVKN 1013
QY 995 LGST-----CHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIPFFGIOEFPNATLK 1049
DB 1014 LTKTDHTECTQSEPCGSDPVQHVKEWHSVVCAL-----TSNK 1053
QY 1050 GNLSPPWYIKTSHNHLIVSTA-----EILFNDVSFTLLPGCAFVRVSRQTEKVPFF-- 1101
DB 1054 ENVTAAEISVGHKTOLRDVSELPILGSEISFNKSLYEGLNAB-----NHRTKITVIFL 1107
QY 1102 --EVPNPLPLVGGSSVGGLLALLALITAAALYKLGFFKRYQKDMSE 1144
DB 1108 KEETRSLPLITGSSIGLLLVVILAILFKCGFFKRYQKDMSE 1152
RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and lymphocytes;
RX MEDLINE=9416496; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shetlerauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";

Genome. Res. 10:165-173 (2000).

[4]

RL MUTAGENESIS OF ASP-109 AND PHE-316.
 RN MEDLINE:20400502; PubMed:10837471.
 RX Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
 RA Brenner M.B., and the role of alpha and beta chains in ligand recognition by beta 7
 RT integrins.";
 RL J. Biol. Chem. 275:25652-25664 (2000).

CC -I- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
 CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
 CC CELL MONOLAYERS.
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
 CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS.
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -I- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -I- SIMILARITY: Contains 1 VWFA domain.
 CC -I- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; L5851; AAB59359.2;
 DR EMBL; AF168787; AAF3107.1;
 DR PIR; A53213; A53213.
 DR HSSP; P11215; IABX.
 DR Genew; HGNC:6147; ITGAE.
 DR MIM; 604682;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 3.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PSS0234; VWFA; 1.
 DR Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 DR Signal; Repeat; Polymorphism; Magnesium; Calcium.

FT SIGNAL 1 18
 FT CHAIN 19 1179 INTEGRIN ALPHA-E
 FT CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.
 FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
 FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1125 1147 POTENTIAL.
 FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 181 198 GLU-RICH (ACIDIC).
 FT REPEAT ? ?
 FT REPEAT ? ? FG-GAP 1.
 FT DOMAIN ? ? FG-GAP 2.
 FT DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).
 FT DOMAIN 200 391 VWFA.
 FT REPEAT 401 456 FG-GAP 3.
 FT REPEAT 457 506 FG-GAP 4.
 FT REPEAT 510 571 FG-GAP 5.
 FT REPEAT 573 638 FG-GAP 6.
 FT REPEAT 641 693 FG-GAP 7.

FT CA_BIND 522 530 POTENTIAL.
 FT CA_BIND 586 594 POTENTIAL.
 FT CA_BIND 654 662 POTENTIAL.
 FT SITE 1150 1154 GPFKR MOTIF.
 FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 126 159 BY SIMILARITY.
 FT DISULFID 706 762 BY SIMILARITY.
 FT DISULFID 823 829 BY SIMILARITY.
 FT DISULFID 893 907 BY SIMILARITY.
 FT DISULFID 1008 1033 BY SIMILARITY.
 FT DISULFID 1041 1057 BY SIMILARITY.
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 360 360 D -> E.
 FT VARIANT 1041 1041 /FTID=VAR_008884.
 FT VARIANT 1041 1041 C -> S.
 FT MUTAGEN 208 208 /FTID=VAR_008885.
 FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.
 FT CONFLICT 477 477 V -> I (IN REF. 3).
 FT CONFLICT 482 482 Q -> R (IN REF. 3).
 FT CONFLICT 950 950 R -> W (IN REF. 3).
 FT CONFLICT 1019 1019 A -> V (IN REF. 3).
 SQ SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;

Query Match 19.5%; Score 1163; DB 1; Length 1179;
 Best Local Similarity 28.8%; Pred. No. 1.3e-70;
 Matches 356; Conservative 221; Mismatches 481; Indels 180; Gaps 41;

QY 5 VLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQG-----SRVVGVAQEIYVAHQ 59
 DB 7 LLCIASLALLAAFNVDVAPMLTPKGGAPFVLSLLHQDPSTNQWLLVTSR-----TKRT 63
 QY 60 RGSLYQDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPQLACGPTVH 110
 DB 64 PGLHRCSLVQDEILCHPVEHVPIPKGRHGRVTVVRSHHGLICI-----QVLVRRP--H 116
 QY 111 QTCSENTYVKGCLFLFGSNLRQPOQ----- 135
 DB 117 SLSSSLT--GTCSLGLPDLRPAQANFFDLENLDPDARVDTGDCYSNKEGGEDDVNT 173
 QY 136 -KFPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL- 180
 DB 174 ARQRALEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEE 233
 QY 181 -KKSKTFLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRHTHTATGIRKVVRELF 239
 DB 234 EKCPFCNFAVQYGGVIOTEFDLRDSQDVMSLARVQNIQVGSVTKTASAMQHLDSIF 293
 QY 240 NITNGARKNAFKILVITDGEKFGDPLGYEDVIPLEADREGVIRYVIGVCDARFSEKSR 299
 DB 294 TSSHGSRRKASKVMVVLTDGGIFEDPLNLTTVINSPKMGQVERFAIGVEEFKARTARE 353
 QY 300 LNTIAKPPDRDHVFOVNNFELKTIQNLREKIFAIEGTQTCSSSSSFEHMEQEGFSAAI 359
 DB 354 LNLIASDDETHAFKVTYNYMALDGLLSKLRNIIISMEGT---VGDALYQLAQIGFSAOI 410
 QY 360 TSNGP-LLSTVGSYDMAGGVFLY-TSEKSTFTINMTRVDSMDNDA-----YLGAAAIILR 413
 DB 411 LDERQVLLGAVGAFDMSGGALLYDTRSRGRFLNQTAAAAADAAEAQAQSYLGAVAVLHK 470
 QY 414 NRVQSLVLGAPRYQHIGLVAMPR-QNTGWESNANV-KGTQIGAVFGASLCSVDVDSNGS 471
 DB 471 TCLSLYVAGAPQYKHG--AVPELOKEGREASFLPVLEGEQMGSYFGSELCPVDIMDGS 528

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF109681; AA01258.1; -
CC EMBL; AF137378; AA051919.2; -
CC EMBL; AL359064; CAB94392.1; -
CC HSP; P17301; IAOX.
CC Genew; HGNC:6136; ITGAL1.
CC MIM; 604789; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC GO; GO:000518; F:collagen binding activity; TAS.
CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
CC GO; GO:0007517; P:muscle development; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; P01185; INTEGRIN.
CC PRINTS; P00453; VWFA DOMAIN.
CC SMART; SM00191; Int_alpha; 5.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
CC PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 1189
FT DOMAIN 23 1142
FT TRANSMEM 1143 1165
FT DOMAIN 1166 1189
FT REPEAT 38 94
FT REPEAT 102 163
FT DOMAIN 167 345
FT REPEAT 359 420
FT REPEAT 422 475
FT REPEAT 477 537
FT REPEAT 539 598
FT DOMAIN 1154 1162
FT DOMAIN 1174 1177
FT CA_BIND 488 496
FT CA_BIND 551 559
FT CA_BIND 613 621
FT DISULFID 76 83
FT DISULFID 121 139
FT DISULFID 129 159
FT DISULFID 659 668
FT DISULFID 674 729
FT DISULFID 781 787
FT DISULFID 881 893
FT CARBOHYD 82 82
FT CARBOHYD 95 95
FT CARBOHYD 291 291
FT CARBOHYD 331 331
FT CARBOHYD 358 358
FT CARBOHYD 449 449
FT CARBOHYD 462 462
FT CARBOHYD 528 528
FT CARBOHYD 642 642
FT CARBOHYD 694 694
FT CARBOHYD 857 857
FT CARBOHYD 894 894

FT CARBOHYD 973 973 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1032 1032 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1040 1040 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 433 433 V -> M.
FT /FTID=VAR_009889.
FT VARIANT 524 524 R -> L.
FT /FTID=VAR_009890.
FT VARIANT 972 972 L -> P.
FT /FTID=VAR_009891.
FT VARIANT 1003 1003 I -> M.
FT /FTID=VAR_009892.
FT VARIANT 1030 1030 MISSING.
FT /FTID=VAR_009893.
FT VARIANT 1094 1094 L -> V.
FT /FTID=VAR_009894.
SQ SEQUENCE 1189 AA; 133609 MW; 60303C08A4A4CD52 CRC64;

Query Match 18.5%; Score 1102.5; DB 1; Length 1189;
Best Local Similarity 28.5%; Pred. No. 1.6e-66;
Matches 354; Conservative 215; Mismatches 503; Indels 171; Gaps 48;

QY 10 ALTLCHG----FNLDTENAMTFQENARG--FGOSVVO--LQSGR-VVVGAPQEIIVAAVQNG 61
DB 12 ALSLWFGFTDTFNMTRKPRVPGSTAFFGVTVQOHDISGNKWLVGAPLENGYQKYG 71
QY 62 SLYQCDYSTGSCPEIRL-----QVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCS 115
DB 72 DVYKCPVIHGNCCTKLNLRVTLSNVSEKDNMRLGLSLATNPKNDSFLACPLWSHECGS 131
QY 116 NTYVKGCLFLPGSNLRQOKPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVST 175
DB 132 SYTTGMCRSVNSNFRFSKTAP-ALQRC-QTYMDIVIVLDGNSNYP--WWEVQHFLLN 187
QY 176 VMBQ--LKKSKTLFSLMOYSEFRHFTFKFONNPRSLVKPITQLLG-RHTHTATGIR 232
DB 188 ILKKFYIGCQIQGVVQYGEDVVFHFLNDRSVKDVVEAASHIEQRCGTERTAFGIE 247
QY 233 KVVRELFNITGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGV----- 287
DB 248 FARSEAFQ--KGGKKGAKKVMIVITDGESHDSF-DLEKVIQOESRDNVTRYAVAVLGYN 304
QY 288 GDAFRSEKSRQELNTIASKPPRQHVQVNNFEALKTIQNLREKIFAIEGTQTGSSSFE 347
DB 305 RRGINPETFLNEIKYIASDPDDKHFFNVTDAAKIDIVDALGRIFSELEGTNK-NETSPG 363
QY 348 HEMSGGSAITNSGPLLSTVGSYDWAGVFLYTSKEK-----STFINNTRVDSNDMDA 402
DB 364 LEMSGTGFSSHVEDGVLLGAVGAYDNNGAVLKETSAGKVIPLRESYLKEFFPELKHGA 423
QY 403 YLGYAAAAILLRNRV-QSLVLGAPRYQHIGLVAMF-RQNTGMWESNANVKQTIGAYFGAS 460
DB 424 YLGYTVTSVSSRQGRVYVAGAPFNHTGKVLFTMHNRSRLTIHQMRGQQIGSYFGSE 483
QY 461 LCSVDVDSNGSTDLVLIGAPHYYEQTR-GGQVSVCPUPRCQARWQCDVLYGEOGQPMG 519
DB 484 ITSVDIDGCVTDVLLVGAPMYFNEGRGKVVYEL---RQRVYVYNGTLKDSHYSQNA 540
QY 520 RFGAALTVLGDVNGDKLTQVAICAPGEEDNRGAVYLFHGTSGSGISPSHSHQAGSKLSP 579
DB 541 RFGSSATVRDLNQDSYNDVVVGAPUEDNHAGAIYIFHGPGRGS-ILTKPQRITASLAT 599
QY 580 RLQVFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQPLVRVKAIMEFNFREAVRVF-- 637
DB 600 GLQYFGCSIHGQLDLNEDGLIDLAVGALGNVILWSRPVQINASLHFESKLI-NIFHR 657
QY 638 ECNDQVVKGEAEVRVCL-----HVQKSTRDLRREGQIQSVVYDLDALDSG 684
DB 658 DC-----KSRGRDATCLAAFLCFTPIFLAPHFQTTTVG-----IRYNATMDER 700
QY 685 RPHSAVFNETKNS-TRRQTVLGLTQCTETLQLPNCIEDPVSPIVLRNLSLVCTPL 743
DB 701 RYTPRAHLDEGGDRFTNRVALLSSGQELCERINFHVLDTADYVVKPVTFVSVEYSLDP-- 757

Db 73 PVDLSTTTCKLNQSTSMNSVTEMTKNNSLGLTLTRNVGTGGFLTCGLPMAQQCGSQY 132
Qy 118 YVUGLCLFGLNLRQOPQKPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRMKFEFVSTVM 177
Db 133 YTTGVCSVSPDF-QLRTSAPAVQICP-SFIDVVVCDSESNIIYWD--AVGNFLEKVF 188
Qy 178 EQLK--KSKTLFSLMQYSEBRFHFTFKFQNNPNRSLVKPITOLL----GRTHATGI 231
Db 189 QGLDIGTPTOMGLIQYANNPRVFNLTFSKD---EMIKATSQTFQYGGDLTNTFKAI 245
Qy 232 KVVRELFTNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYIVG--- 287
Db 246 QYARDTAYTAAGRGPATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAGLVYL 304
Qy 288 -GDAFRSEKSRQELNTTASPPRDRHVFQVNNFEALKTIONQLREKIFAIGTGTGSSSF 346
Db 305 NRNALDTNLIKELIKATASPTERHPFNVSDEADLLEKAGTIGEIPFISGTVQG-GDNF 363
Qy 347 EHEMSQBFSAAIT--SNGPLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-D 401
Db 364 QMEMSQVGFSAEYSPQNNILMLGAVGAYDWSGTWQKTPHGLIFSKQAPEQILQDRNHS 423
Qy 402 AYLGAAAILNRVQSLVIGAPRYQHIGLVAMPFRONTGHWESNAV-----KGTQIGA 455
Db 424 SYLGYSVASISTGNSHFVAGAPRANYTGQILVLSVN-----ENGNVTVIQSRGDDIGS 478
Qy 456 YFGASLCSVDVDSNGSTDVLIGAPHYVEOTR--GGQVSVCPPLRGQRARWQCDVAVLYGE 513
Db 479 YFGSVLCAVDVNDKTDITDVLVIGAPMYNDLKKEEGRVYLTITKG-ILNWH--QFLEGP 535
Qy 514 QGQPMWFGAALTVLGVNDGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIA 573
Db 536 NGLNARFGSAIALSDINMDGFNDVIVGSPLENQNSGAVVYNGHEGM-IRLRYSQKIL 594
Qy 574 GS--KLSPRLQYFCQISGGODLTMDGLVLTUACQGHVLLRQOPVLRYKATMEFPRE 631
Db 595 GSDRAFSSHLYQFGRSLDGYDLNGDSITDVSQVAFQGVQVQLWSQSIADVSVDASFPPK 654
Qy 632 VARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQISVVYTDLALD----SGRPH 687
Db 655 I--TLNKNAEI-----KLKLCF---SAKFRPTQNNQVAIVNITDEDQFSRVI 701
Qy 688 SRAVNETKSTRQTVLGHQTQCE--TLKLQPLNCIEDPVSPIVLRNPSL--VOTPL 743
Db 702 SRGLFKENNERCLQKTMIVSQARCSEVYIIHQPS---DIISPLNLCMNISLENPGT-- 756
Qy 744 SAFGNLRVLAEADQRLFTALFPPEKNCNDNIQDDLSITF----SFMSLDCIACVGGPR 799
Db 757 -----NPALRAYSETVXFSIPPHKOCGDDGVCSIDLVLNVQQLPATQOQPFIVSNQK 810
Qy 800 EFNVTVTVRNDEGDSYRTOVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSG 858
Db 811 RLTSVQLKNKEAYNTEIVDSENLFF-----ASWSPVDGTEVTCQIAS 858
Qy 859 ALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTBFQL 918
Db 859 SOKSVTCNVGYPALKSKQVTFITNFDNLQ-NLQNOASISFRALSSEOBENMADNSVNL 917
Qy 919 ELPKYAVYVMTVSHGVSTKVLNTFASENTSRVQHQVQSNLQOR-----SLPI 968
Db 918 KLSLLYDAEIHIT-RSTNINFYVSLDGNVSVV-HSFE--DIGPKFISIKVTGVSVPV 973
Qy 969 SLVFLVRLNQTWIDRPOVTFSEN---LSSTCHTKE-----RLPSHS 1009
Db 974 SMA-----SVIIHPQTOKKNPLMYLTGHTDQAGDISCEAEINPLIKGTSSSV 1024
Qy 1010 DFLAE-LRKAPVNCISAVCORIQCDIPFGIQEEFNATLKNLSFDWYIKTSHNHLIV 1068
Db 1025 SFKSENPRHKLNCRTASCNIMCWLRLDQVKGVEFLNVSTRIWNGTFAASTFTQVQLT 1084
Qy 1069 STAEI-LFNDSVFTL-----LPGQAFVRQSTETKVEPFE-VNPLPLIVGSSVGILL 1120
Db 1085 AAAEIDTNPQIVYIENVTITP-----LTIMPKHEKVEPTGTIVGSVIAGILL 1134

Qy 1121 LALITAALYKLGFFKQYKQYKDM 1141
Db 1135 LLALVAILLWKLGGFFKRYKYM 1155

RESULT 12

ITA2_MOUSE
ID ITA2_MOUSE STANDARD; PRT: 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding."
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development."
RL Dev. Dyn. 199:332-314(1994).
CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29987; CAA82877.1; -;
DR EMBL; X75427; CAA53178.1; -;
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR004113; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; P800242; INTEGRIN_ALPHA; 1.
 DR PROSITE; P850234; VWA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 26
 FT CHAIN 27 1178
 FT DOMAIN 27 1129
 FT TRANSMEM 1130 1151
 FT DOMAIN 1152 1178
 FT REPEAT 1 100
 FT REPEAT 2 3
 FT DOMAIN 185 375
 FT REPEAT 3 483
 FT REPEAT 485 546
 FT REPEAT 548 607
 FT REPEAT 612 664
 FT CA_BIND 496 504
 FT CA_BIND 560 568
 FT CA_BIND 624 632
 FT SITE 480 482
 FT SITE 1154 1158
 FT DISULFID 80 89
 FT DISULFID 677 734
 FT DISULFID 786 792
 FT DISULFID 862 873
 FT DISULFID 1016 1047
 FT DISULFID 1052 1057
 FT CARBOHYD 102 102
 FT CARBOHYD 109 109
 FT CARBOHYD 429 429
 FT CARBOHYD 457 457
 FT CARBOHYD 472 472
 FT CARBOHYD 696 696
 FT CARBOHYD 1054 1054
 FT CARBOHYD 1071 1071
 FT CARBOHYD 1078 1078
 SQ SEQUENCE 1178 AA; 128926 MW; 1f194B9C0240F465 CRC64;

Query Match 18.3%; Score 1089; DB 1; Length 1178;
 Best Local Similarity 28.1%; Pred. No. 1.3e-65;
 Matches 348; Conservative 209; Mismatches 492; Indels 188; Gaps 44;

QY 5 VLLLTALTLCGPNLDTENAMTFQ--ENARGFGOSVVLQ---QGSVVVVGAPQEIIVANQR 60
 DB 15 LMLVQQLNCLATNVLGRLPAKIFSGPSSQFGVSVQVLTNPQGNWLLVSGSPWGFENRM 74
 QY 61 GSLYQC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQ 111
 DB 75 GDVYKCPVDLPATACEKINLONSASISNVTETKNMSLGLTLTRNPGTGGFLTCGLPLWAH 134
 QY 112 TCGENTYVKGCLPFGSNLRQQPQ---KPEALRGCPQEDSDIAFLIDGSGSIIPHDPRR 168
 DB 135 QCGNQYVATGIC---SDVSPDFQLTSPFAVQACPSL--VDVVVCDENSIYIP--WEA 187
 QY 169 MKEFVSTWMEQLK--KSKTLFSLMOYSEERIHFTKEFQNNPNPSLVKPIQLLG--RT 225
 DB 188 VKNFLVKFVTGLDIGPKTKQVALIOANEPRIIFNLNDFETKEDMVOQATSETRQHGDDLT 247
 QY 226 HTATGIRKVVRELFNITNGARKNAFKILVVITDGEKFGDPLGVEDYIPEADREGVRYVI 285
 DB 248 NTRALEPARDAYAYSOTSGRPGATKVMVVVTDGESH--DGSKLKTIVIQCNDDIILRFGI 306
 QY 286 GV-----GDAFSEKSRQELNTIASPPRDHVFQVNNFALKTIQNLREKIPAIETQT 340
 DB 307 AVLGLNRLNADTKNLKEIKAIATSTPTERYFPNVADAALEKAGTGLGEQIFSIETGTQ 366
 QY 341 GSSSSFEHMSQGFSA--AITSNGPLSTVGSYDWAAGVFLYTSKEKSTFFINMT--RVD 396
 DB 367 G-GDNFMQMAQVGFSGADYAPQNDIILMLGAVGAFDWSGLTVQETSHKPIVFPKQAFDQVL 425

QY 397 SDMN--DAYGYAAAIIILNRVQSLVLAGPRVQHIGLVAMFRONTQGWESNANV-----KGT 451
 DB 426 QDRNHSFGLGSVAIAISTEDGVHFVAGAPRANVTQCIVLYSYK---QGNVTVIQSHRGD 482
 QY 452 QIGAYFGASLCSVDVDSNGSDLVLIIGAPHYVEQTR--GGQVSVCPPLRGGQRRWQCDV 509
 DB 483 QIGSFVGSVLCSDVDKDTITDVLVVGAPTYWDLKKEEGVYLTFTIKGLINQIQ---F 539
 QY 510 LYGEQGPWGRFGAALTVLGDVNGDKLTVAIGAPEGEDNRGAVLPHGTSGSISPSHS 569
 DB 540 LEGPEGTGNARFGSAIALSDINMDGPDVIVGSPVENENSGAVVIYNGHQT--IRTKYS 598
 QY 570 QRIAGSKLSPR--IQYFGOSLSCGGDLTMDGLVLTVAQAGHVLRLRQPVLRVKAIMEF 627
 DB 599 QKILGNGAFRRHLQFFGRSLDGYDGLNGDSITDVSIGALGQVQLWQSQSIADVAIEALF 658
 QY 628 NPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALO---S 683
 DB 659 TP-----DKITLLNKDAKITLCLCFRAEPFAGQNNQV---ALLFNMTLQDAGHS 705
 QY 684 GRPHSRVFNETKNSRRQTVLGLTQTCET--LKLQPLNCIEDPVSPIVLRNLSLVGT 741
 DB 706 SRVTSRGVFRNSERFLQKNMVNEVQKSEHHISIQKPS---DVNPLDLRAVDISLENP 762
 QY 742 PLUSAFGNLRPVLAEDAQRLFTALPFPEKNCNGNDICQDDLSI-----TFSPMSLDC 792
 DB 763 GTS-----PALEYSETVKVFSIPFYKECGSDGICISDLILDVQQLPAIQTSF----- 811
 QY 793 LVVGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRSLACESAS 852
 DB 812 IVSNQNKRLTFSVLNKGESAYNTVLAEFSENLF-----ASFMPVDGTE 859
 QY 853 ST-EVSGALKSTSCSINHPIFPENSEVFNTIFDVDSKASLGNKLLKANVTSENMMPT 911
 DB 860 VTCEVSSQKSVTCVGVPAKSEQOVTTFINDFNLQ--NLQQAIAINFOAFSESO--ET 916
 QY 912 NKTE--FOLELPVKYAVVMVTSHGVSTKYNLFTASENSTRVMQHOYQVSNLQOR----- 964
 DB 917 NKADNSVSLTIPLLYDAELHLT-RSTNINFVEISSDENAPSVIK---SVEDIGPKFIFSL 972
 QY 965 -----SLPISLVFLV-----PVLNQTVIWD 985
 DB 973 KVTAGSAPVSMALVTHIPIQYTKENPLLYLTGIDTQDQAGDISCTAEINPLKPHTA--- 1029
 QY 986 RPQVTF--SENLSSTCHTKERLPSPSHDFAELARKAPVWNCIAVCORICQDIPFFGQIEEF 1044
 DB 1030 -PSVGFKNFR--HTKE-----LDCRTTSCSNITCWLKDLHMKAEY 1068
 QY 1045 NATLKNLSFDWYIKTSHNHLIVTAEILFNDVSFTLLPGQCAFVRSQTETKVEPEVP 1104
 DB 1069 FINVTRVMNRFAASTFTQVLTAAAEIDTHNPQLFVIEENAVTIPLMKPTEKAEVP 1128
 QY 1105 NPLPLIVGSSVCGILLALLITAALYKLGFFKQYKDM 1141
 DB 1129 T--GVIIGSIAGILLAMTAGLWKLGGFFKQYKDM 1163

RESULT 13
 ITA2_HUMAN
 ID ITA2_HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN ITGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.

FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT TURN 313 317
FT TURN 318 319
FT TURN 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 18.3%; Score 1087.5; DB 1; Length 1181;
Best Local Similarity 27.3%; Pred. No. 1.7e-65;
Matches 339; Conservative 214; Mismatches 498; Indels 193; Gaps 44;

QY 3 LRVLTLTALT-----LCHGFNDLTENAMTFQ-ENARGFGQSIVQL--QGSRVVVGAPQE 53
DB 11 LPLLVLALSQILNCCLAAYNGLPEAKIFGSPSEQFGYAVQQPINPKGNWLLVGSWS 70

QY 54 IVAANQSGSLYQC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPQLLA 104
DB 71 GPFENRMGDVYKCPVDLSTACEKLNLOTSTISIPNVTETMKNMSLGLILTRNMGTFGFLT 130

QY 105 CGPTVHQTCSENTYKGLCFPLGSLNRLOQPOKFPALRGCPQEDSDIAFLIDGSGSIIPH 164
DB 131 CGPLWAQCGNOYTTGVCSDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPW 188

QY 165 DFRMRKFRVSTVMEOLK--KSKTLFSLMOYSEEFRIHETFEFQNNPNRSLVPIKTCOLL 222
DB 189 D--AVKNLEKFEVQGLDITGPTKTQVGLIQYANNPRVFNLTNYTKEEMIVATSOYSG 246

QY 223 G-RHTATGIRKVRRELFINIENGAKNAFKILVITDGEKFDPLGVEDVPEADREGVI 281
DB 247 GDLTWTFGAIQYARKVAYSAAGRRSATKVMVVTDESH-DGSMKAVIDQCHNDIL 305

QY 282 RYVIGV-----GDAPFRSKROELNTIASKPRDHVQVNNPEALKTIONLREKIPAI 336
DB 306 REGIAGVGLNRLNADLTNLIKAIKIASIPTEYFFNVSDAALLERKAGTLGEQIFSI 365

QY 337 GTQTCSSSFHEMSQEGSAITSNGP--LLSTVGSVDWAGGVFLYTSKEKSTFINWT- 393
DB 366 GTVQG-GDNFQEMSQVGSADYSSQNDILMLGAVGAFQWGSITIVQKTSHGHLIFPKQAF 424

QY 394 -RVDSMDN-DAYLGAIAAILRNRYQSLVGLAPRYCHIGLVAMFRONTGMWESNANV-- 448
DB 425 DQILQDRNHSSYLGYSVAISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNI 479

QY 449 ---KGTQIGAYFGASLCSYVDVDSNGSTDVLILGAPHYVEQTR--GGQVSVCLPRGOR 503
DB 480 QAHRGDQIGSYFGSVLCSYVDVKDITDVLVVGAPWMSDLKKEGRVYVLFITKKGILGQ 539

QY 504 WQCDVILGEOQCPNGRFGAALTVLGDVNGDKLTVAIGAPCEENRGAIVLPHFTSSG 563
DB 540 HQ---FLEGPEGIENTRFGSAIALSNDIMGDFNDVIVGSPLENQNSGAVIYHGHQT- 595

QY 564 ISPSHSQRTAGS--KLSPLQYFGOSLSGGDLTDWGLVDLTVGAGHVLLRLRSOPVLVR 621
DB 596 INTKYSQKILGSDGAFRSHLQYFGRSLOGYDGLNGDSITDVSIGAFQGVQVQWLSQSIAD 655

QY 622 KAIMEFNPREVARNYFECNDQVVGKEAGEVRVCLHVQKSTRDLREGIQISVWYTDIAL 681
DB 656 ALEASFTEPEKI--TLVNKNAQII-----LKLCP-----SAKFRPTKQNNQVAIVYNTL 702

QY 682 D-----SGRPHSAVFNETKSTRQTVLGLTQTC--ETLKLQLENCDIEDPVPVILRLN 735
DB 703 DADGFSRVRTSRLGFKENNERCLQKNMVVNOAQSCPEHIIYIQEPS---DVVNSLDLRVD 759

QY 736 FSLVGTPLSAGNLRPVLAEDAQRLLFTALPFFKXKNCNDNICQDLSITF-----SPMSLD 791
DB 760 ISLENPGTS-----PALEAYSETAKVFSIPFFHKDCGEDGLCTSDLVLRQIPAAQEQP 813

QY 792 CLVVGGRPRENVTVTVNRDGEDSYRTOVTFPFFPLDLSYRKVSTLQNRQSRQSLACESA 851
DB 814 FIVSNQKRLTFSVTLKKRESAYNTGIVDFSENLF-----ASFSLPVDGT 861

QY 852 SST-EVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPR 910
DB 862 EVTCQVAASQKSVACDVGYPAKREQQVTFTFINFDNLQ-NLQNAQSLSQALSESQEN 920

QY 911 TNKTEFOLELPVKYAVYVMTVSHCVSTKYLNFNTASENTSRVMOHQYQVSNLQOR----- 964
DB 921 KADNLVNLKIPLLYDAEI---HLTRSTNINFEISSDGNVPSTIVSHFEDVGPKFIFSLK 976

QY 965 ---SLPISLVFLV-----PVRLNOTVIWDR 986
DB 977 VTTGSPVSPVATVIHIIPOYTKENKPLMYLTGVOTDKAGDISCNADINPLKIGOT-----S 1032

QY 987 PQVTF-SENLSSTCHTKERLPSSHDFLAELRKAPVNVNCISAVCQRIQCIDIPFFGIQEEFN 1045
DB 1033 SSVSFKSENER---HTKE-----LNCRTASCNSVTCMLKDVHMKGEYF 1072

QY 1046 ATLKNLSFDWYIKTSHNHLIVSTAEI-LFNDVSFTLLPQGGAFVRSQTTETKVEPEVP 1104
DB 1073 VNVTTIRWNGTFASSTFQVLTAAAEINTVPEIYVI-----EDNTVTIPLMIM 1122

QY 1105 NP-----LP--LIVGSSVGGLLLLALITAAALYKLGFFKRYKDM 1141
DB 1123 KPDEKAEVPTGVTIIGSIAGILLALLVAILKLGFFKRYKYM 1166

RESULT 14
ITAG_HUMAN STANDARD; PRT; 1167 AA.
AC 075578; Q9UHZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Aakerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
RT alpha10, a beta1-associated collagen binding integrin expressed on
RT chondrocytes.";
RL J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
RA Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
RT structure, and chromosomal localization.";
RL Cytogenet. Cell Genet. 87:238-244(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN
CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.


```
QY 92 SENLSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIPFFGIQIEBFAFNATLKN 1051
DB 1009 TNNASCIQVNLTEPPCPVPHPELOHTNRLNGSNSTQCQVVRCHLQGLAKGTSEVSGLLRL 1068
QY 1052 LSFDFWIKTSHNHLIVSTAEILFVDSVFTLLPGQAFVRSOTETKVPPEFVNPPLIV 1111
DB 1069 VNEEFRAKPKSLVSTFELGTGEGVQLQTEASRWSESLEEV-VQTRPILISLWILI 1127
QY 1112 GSVGGLLLALITALYKLGFF-----KQYK 1139
DB 1128 GSVLGGLLALLVCLWKLGFPAHKKIPEBKREK 1164

RESULT 15
ITAL RAT STANDARD; PRT; 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGA1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch F., Carbonetto S., Reichardt L.F.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
RT for laminin and collagen."
RL J. Cell Biol. 111:703-720(1990).
RN [2].
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Kotliansky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alphabeta1 integrin I-domain: insights into
RT integrin I-domain function";
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL X52140; CAA36384.1;
DB PIR A35854; A35854.
DB PDB 1CK4; 03-MAY-00.
DB InterPro: IPR000413; Integrin_alpha.
DB InterPro: IPR002035; VWF A.
DB Pfam PF01839; FG-GAP; 3.
DB Pfam PF00357; integrin_A; 1.
DB Pfam PF00092; vwa; 1.
DB PRINTS: PRO1185; INTEGRINA.
DB PRINTS: PRO0453; VWFADOMAIN.
```

```
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1180 INTEGRIN_ALPHA-1.
FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1143 1165 POTENTIAL.
FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 175 388 VWFA.
FT REPEAT 377 432 FG-GAP 3.
FT REPEAT 433 484 FG-GAP 4.
FT REPEAT 485 565 FG-GAP 5.
FT REPEAT 567 626 FG-GAP 6.
FT REPEAT 629 681 FG-GAP 7.
FT CA_BIND 497 505 POTENTIAL.
FT CA_BIND 579 587 POTENTIAL.
FT CA_BIND 641 649 POTENTIAL.
FT SITE 1168 1172 GFGR MOTIF.
FT DISULFID 82 92 BY SIMILARITY.
FT DISULFID 687 696 BY SIMILARITY.
FT DISULFID 702 755 BY SIMILARITY.
FT DISULFID 807 813 BY SIMILARITY.
FT DISULFID 877 885 BY SIMILARITY.
FT DISULFID 1029 1062 BY SIMILARITY.
FT DISULFID 1066 1073 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1180 AA; 130808 MW; 855DA2BE02362EE4 CRC64;
```

Query Match 18.2%; Score 1082; DB 1; Length 1180;

Best Local Similarity 27.7%; Pred No. 4e-65;

Matches 351; Conservative 199; Mismatches 484; Indels 234; Gaps 47;

QY 6 LLLTALTLCGHNLDENAMTFQGVVQL---QGSRYVVGAPQEIIVAAHQ 61

DB 18 LLTVLGFCSFNVVDVKMSFSFGPVEDMFQYVQYENEGKWLIGSLVGPQKARTG 77

QY 62 SLYQCDYSTGSCPE-IRLQVPVEA-----VMSLGLSLAATSPQLLACQPTVHQ 111

DB 78 DVYKCPVGRERAMPCKVLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGFLACPLVAY 136

QY 112 TCSENTYVYKGLCFGLGSLNRQOPKQPEALGCPDESDIAFLIDGSGSIIPHDFRMKE 171

DB 137 RGHLYHTTGICSDVSPFTFQVNSFAP--VQECSTQ-LDIIVLDGSGSIIP--WESVIA 191

QY 172 FVSTVMEQLK--KSKTLFLSMLQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLGR-RHTA 228

Db 192 FLNDLLKMDIGPKQTVQVIGVQGENVTHEFNLYKSSSTEVEVLVAANKIGROGLQMTA 251
Qy 229 TGRKVVRELFNITNGARONAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGV 288
Db 252 LGIDTARKEAFTARGARRGVKVMWIVTDGESH-DNYRLKQVIQCEDENIQFSAIAL 310
Qy 289 DAFR-----SEKSRQELNTIASPPRDHVFQVNFALKTIONQLREKIPAEGTGTGSS 343
Db 311 GHYNRGNLSKEKFEVEIKSIASEPTEKHFNVSDALALVTIVKALGERIFALEATADQSA 370
Qy 344 SFEHEMSQGFSAAITSGNPLLSITVGSYDWAGVFLYTSKESKSTFINMT--RVDSDMD 401
Db 371 ASPEMENSOTGFSAHYSQDWMILGAYDWMGTVMQKANOVIPIHNTTFFOTEPAKMNE 430
Qy 402 ---AYLGYAAAIILNRNVQSLVLCAPRYOHIGLVAMPONTGMWESNANVKTQIGAYFG 458
Db 431 PLASYLGYTNSATIPGDVLYIAGOPRYNHTQVVIYKMEDGNINILQTLGGEQIGSYFG 490
Qy 459 ASLCSVDVDSNGSTDVLYIGAPHY-----YEQTR-GGQVSVCPPLPRQRA 502
Db 491 SVLTIDIDKDSYTDLLVAGPMYMGTEKEEQGVVYVAVNQTRFEYQMSLEPIROTCCS 550
Qy 503 RWQCDVLYGEOGQPWG-REGAALTVLGDVNGDKLTDVAIGAPGEEDNRCANVLFHCTSG 561
Db 551 SLKDNSCTKENKNEPCGAREFGTAIAAVKDLNVDFNDWIGAPLEDDHAGAVIYHG-SG 609
Qy 562 SGISPSHSORLAGSKLPRLOQFGOSLGGDLTMDGLDVLTVGAQGHVLLLRSQPVLV 621
Db 610 KTIREAYAQRIPSGDKTKLFFQOSIHGEMDLNGDGLTDTVTGGGLGAALFWARDVAV 669
Qy 622 KAIMFNPREVARNVPCNDQVKGKAG--EVRVCLHVQ-KSTRDLREGQIQSVVYTD 678
Db 670 KVTWNPFPKVIQKKCR---VEGKETVCINATMCPHVKLKSKEDSIYEADLQ----YR 722
Qy 679 LALDSGRPHSRVFNET-----KNSTRQTVQLGLTQTCETLKLQLPNCI----- 723
Db 723 VTLSLRQISRSFSGEQERKIQRNIITRESE-----CIRHSFYMLDK 765
Qy 724 EDPVSPIVLRNFSLVGTPLSAGFNLRPLAEDAQRULTALFPFEKNCNGNDNICQDDLSI 783
Db 766 HDPQDSVRVLDFNLT-DPENG-----PVLDDALPNSVHEHIPFAKOCGNKERCISDITL 819
Qy 784 TFSFMSLDCLVVGPRE-FNVTVTRNDGEDSYRTVTFPPFDLSYRKVSTLQNRQSR 842
Db 820 NVSTTEKSLLVKSHDKFNVLTVKNKGSAYNTRTVQHSNLI FSGIEEIQK--- 875
Qy 843 SWRLACSASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLL-LKAN 901
Db 876 ----SCESN-----QNTICRVGYPFLRAGETVTFKIIQFNTSHLSENAIHLSAT 922
Qy 902 VTSENMPRTNKTFFOLEPVKYAV---YMWVTSHGVS-----KYLNFASENTSR 950
Db 923 SDSEEPLESNDNEVNSIPVYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGN 980
Qy 951 VMHQYQVSNLQORSLP---ISLVF-----LVPVRLNQTVIWD-----RP--- 987
Db 981 EINVFYITIRKRGHPMPPELQLSISFPNLTADGYPVLYPIG-----WSSSDNVNCRPSL 1034
Qy 988 -----QVTF-----ENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAV 1027
Db 1035 EDPFGINSCKMTISKSEVLKRGTTIQDCSTC-----GVATITCSLLP 1077
Qy 1028 CORIQCDI-----PFFGIQEEF---NATLGNLSFDWYIKTSHNHLIVSTABILFNDS 1078
Db 1078 SDSLQVNVSLLLMKRTPF-IRAHFSSLNLTLRGELK-----SENSLTLSSN----- 1123
Qy 1079 VFTLLPGQCAFVRSOTETKVPFEPVNPPLP--IVGSSVGGLLLIATAALYKLGFFKR 1136
Db 1124 -----RKRELAIQISKDGLPGRVPLWILLSAFAGLLMLLMLLALMKIGFFKR 1172
Qy 1137 QYKMMSE 1144

Db 1173 PLKKMEK 1180

Search completed: November 25, 2003, 14:17:21
Job time : 13.3033 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.5727 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGHFNLD.....FKROYKDMSEGGPPGAEPQ I153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4459	74.9	1151	11 Q9J130	Q9J130 rattus norv
2	3976.5	66.8	1036	11 Q8CA73	Q8CA73 mus musculu
3	3824.5	64.2	920	6 Q28984	Q28984 sus scrofa
4	3514	59.0	1169	4 Q81VA6	Q81VA6 homo sapien
5	3352.5	56.3	1169	11 Q9QXH4	Q9QXH4 mus musculu
6	3282.5	55.1	1161	11 Q9QYE7	Q9QYE7 rattus norv
7	1545	25.9	1161	11 Q9WTV4	Q9WTV4 mus musculu
8	1534.5	25.8	1160	11 Q9R200	Q9R200 mus musculu
9	1419	23.8	1196	13 Q98TF1	Q98TF1 cyprinus ca
10	1368.5	23.0	1187	13 Q98TF0	Q98TF0 cyprinus ca
11	1361	22.9	1086	4 Q96HB1	Q96HB1 homo sapien
12	1277	21.4	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1171.5	19.7	1167	11 Q88340	Q88340 rattus norv
14	1125.5	18.9	1167	11 Q88341	Q88341 rattus norv
15	1068.5	17.9	1171	13 Q42094	Q42094 gallus gall
16	1049	17.6	1038	11 Q8BS01	Q8BS01 mus musculu

Q8mkf4 felis silve
Q9wuf8 mus sp. itg
Q8tes5 homo sapien
Q8bpq8 halocynthia
Q63001 rattus norv
Q60271 xenopus lae
Q8wy18 homo sapien
Q8ce84 mus musculu
Q61989 mus musculu
Q8c270 mus musculu
Q91vd5 mus musculu
Q8bgu3 bos taurus
Q9ub90 lytechinus
Q76378 lytechinus
Q98tt7 gallus gall
Q9u6sl strongyloce
Q42598 xenopus lae
Q91779 xenopus lae
Q924w2 rattus norv
Q8cc06 mus musculu
Q9mzd6 bos taurus
Q9gk48 bos taurus
Q9tun4 oryctolagus
Q9cun6 sus scrofa
Q70304 mus musculu
Q8sy51 drosophila
Q8bq25 mus musculu
Q9bpq7 halocynthia
Q9tu44 canis famil

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
ID Q9J130
AC Q9J130;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zeria K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF268593; AAF81280.1; .
DR HSSP; F11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.9%; Score 4459; DB 11; Length 1151;

Best Local Similarity 73.5%; Pred. No. 3.3e-317;
Matches 847; Conservative 144; Mismatches 160; Indels 2; Gaps 2;

QY 1 MALRVLLTALTLCGHFNLDTENAMTFENARGFCQSVVQLGSRVVVVCAPQEIIVANOR 60

DB 1 MTLKVLATVLTLCGHFNLDTPNMTFOENASFGQSVIQLGETRVVVAAPQEVKAVNQT 60


```
QY 601 DLTGAGQHVVLLRSQVLRVKAIMEFNPVARNVFECDQVQVKGAGEVRVCLHWOK 660
Db 483 DLAVGAGQHVVLLRSQVLRVKAIMEFNPVARNVFECDQVQVKGAGEVRVCLHWOK 542
QY 661 STDRLEGOISVTVYDIALDSGRPHSRAVNETKNSRROTUVGLTQTCETLKLQIP 720
Db 543 NTKDRUREGDIQSTVTVYDIALDSGRPHSRAVNETKNSRROTUVGLTQTCETLKLQIP 602
QY 721 NCTEDPVSPIVLRNLSVGTPLSAFAGNLRPVLAEADQRLFTALPFPEKNCNDNICQDD 780
Db 603 DCVDDSVSPILRLNLTAGEPLRSFGLNLRPVLANDAQRFATMPFPEKNCNDNICQDD 662
QY 781 LSTFPMNSDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKVSTLQNRS 840
Db 663 LSTFPMNSDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKVSTLQNRS 722
QY 841 QRSWRL-ACESASSSTEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLK 899
Db 723 KKPWFVKPAPRESSSSSGHGLKSTTWNINHPFIPANSEVTEFNITPDVDSHAFGNKLLK 782
QY 900 ANYTSNNPRTNKTKEFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMQHOYQVS 959
Db 783 AIVASENNRTHKTKEFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMQHOYQVS 842
QY 960 NLQORSPLSLVPLVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAP 1019
Db 843 NLQORSPLSVVFWIPVQINNVTMDHPQVTFSENLSSTCHTKERLPSHDSFLAELRKAP 902
QY 1020 VVNCSTAVCQRIQCDIPFGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILLPNSV 1079
Db 903 VVNCSTAVCQRIQCDIPFGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILLPNSV 962
QY 1080 FTLLPQOGAFVRSOTETKVEPPEVNPPLPLIVGSSVGLLALLALITAAALYKLGFFKQYK 1139
Db 963 FALLPQOESVRSKTEKVEPPEVNPPLPLIVGSSVGLLALLALITAGLYKLGFFKQYK 1022
QY 1140 DMMEGGPPGAEPQ 1153
Db 1023 DMNEAAPQDAPQ 1036
RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (T-EMBLrel. 01; Created).
DT 01-FEB-1997 (T-EMBLrel. 02; Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23; Last annotation update)
DE C011b (Fragment).
GN C011b.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9823;
RN [1]
SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR ENBL; U40072; AAB16869.1; -.
DR HSP; P11215; IBIHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-CAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.2%; Score 3824.5; DB 6; Length 920;
Best Local Similarity 79.8%; Pred. No. 6.3e-271;
Matches 735; Conservative 77; Mismatches 108; Indels 1; Gaps 1;
QY 134 POKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKFEVSVMBOLKSKTLESLMOYS 193
Db 1 POKPEALRGCPQEDSDIAFLIDGSGSIINRLOFQRMKEFVSTMVQGFQSKTFLFALMOYS 60
QY 194 BEFRTHFTKEFQNNPNPRSLVKPITQLLGRTHATGIRKVVRELNFNTNGARKNAFKIL 253
Db 61 EDFYTHFTFNDPKRNPSPKLLVRPIQLLGRTHATGIRKVVRELNFHKSAGARENALAIL 120
QY 254 VVITDGEKFGDPLGVEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPRDHVF 313
Db 121 VVITDGEKFGDPLGVEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPRDHVF 180
QY 314 QVNNFEALKTQNLQREKIFALETGTQSSSSFEHEMSQEGFSAITNSGPLLSTVSGSYD 373
Db 181 QVNNFEALKTQNLQREKIFALETGTQSSSSFEHEMSQEGFSAITNSGPLLSTVSGSYD 240
QY 374 WAGGVFLYTSKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVA 433
Db 241 WAGGAFLHMPKDRVIFINTTRVDSMDNDAYLGAAIILNRVQSLVLGAPRYQHIGLVA 300
QY 434 MFRQNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQVSV 493
Db 301 MFKQNSGAWEKNAIDKSGIGSYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQVSV 360
QY 494 CPLPGQARWQCDVAVLGEQOPGRGAALTVLGVNGDKLTQVAIGAGEEDNRGAV 553
Db 361 CPLPGQ-RAKQWQCRVILCEQOPGRGAALTVLGVNGDKLTQVAIGAGEEDNRGAV 419
QY 554 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQCSLSGGQDLTMDGLVLTVCAGQHLVLL 613
Db 420 YLFHGTSELGISTPSHSQRIAGSKLSPRLQYFQCSLSGGQDLTMDGLVLTVCAGQHLVLL 479
QY 614 RSQPLVRKAIEMFNPREVARNVFECDQVQVKGAGEVRVCLHWOKSTRDLREGQIQS 673
Db 480 RSQPLVRKAIEMFNPREVARNVFECDQVQVKGAGEVRVCLHWOKSTRDLREGQIQS 539
QY 674 VVTVYDIALDSGRPHSRAVNETKNSRROTUVGLTQTCETLKLQIPNCIEDPVSPIVLR 733
Db 540 VVTVYDIALDSGRPHSRAVNETKNSRROTUVGLTQTCETLKLQIPNCIEDPVSPIVLR 599
QY 734 LNFSLVGTPLSAFGNLRPVLAEADQRLFTALPFPEKNCNDNICQDDLSITFSFMSLCL 793
Db 600 LNFSLVGTPLSAFGNLRPVLAEADQRLFTALPFPEKNCNDNICQDDLSITFSFMSLCL 659
QY 794 VVGGPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWRLACESASS 853
Db 660 VVGGPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWRLACESASS 719
QY 854 TEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLKANYTSNNPRTNK 913
Db 720 TEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLKANYTSNNPRTNK 779
QY 914 TEFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFL 973
Db 780 TEFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFL 839
QY 974 VPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVWNCSTAVCQRIQ 1033
Db 840 VPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVWNCSTAVCQRIQ 899
QY 1034 DIPFGIQEENATLKNLSF 1054
Db 900 DIPFGIQEENATLKNLSF 920
RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
ID Q81VA6
AC Q81VA6;
```



```
Db 69 CGYHTGNCPEIPISLQVPEAVNISGLSLAAATNPWSWLLACGPTVHHTCRENIYLTGLCLF 128
QY 126 FGSNLRRQKQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFEVSTVMEQLKSKT 185
Db 129 LSSSPKQS-QNFTTAQBCPCQDQDVLFLIDGSGSISSTDFEKNLDFVKAVMSQJQRPT 187
QY 186 LPSLMQYSBEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFINITGA 245
Db 188 RFLSMQFSDFYRVHFTFNNFISTSSPLSLGSRVQLRGVYTTASAKHIVITELFTTQSCA 247
QY 246 RKNAPKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKOELNTIAS 305
Db 248 RODATKVLIVITDGRKQGNLSYSDVIPMAEASIIRYAIGVKAFYNEHSHKQELKAIAS 307
QY 306 KPPRDHVFQVNFPEALKTIQNLREKI FAIEGTQTCSSSFEHEMSQEGFSAITNSGPL 365
Db 308 MPSHEVVSFVNFDAKJENQUKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTDGPV 367
QY 366 LSTVSGYDWAGGVFLYTSKESKTFINMTRVDSMDNDAYLGAAAIILNRVQSLVILGAPR 425
Db 368 LGAVGFSFGAGFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPR 427
QY 426 YQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYEQ 485
Db 428 HQHTGKVIIFTQESRHRPKSEVRGTQIGSYFGASLCSDVMDRDGSTDVLVIGVPHYEH 487
QY 486 TRGGQVSVCLPRGQARQCDALVXGEOGPWREGAALTVLGDVNGDKLTDVAIGARG 545
Db 488 TRGGQVSVCPMP-GVGRWHCGTTLHGEQHPWREGAALTVLGDVNGDSLADVAIGARG 546
QY 546 EEDNRGAVTLFHGTSGSGISPSHSORIASKLSPLRYQFGQSLSGQDGLTMDGLVDLTVG 605
Db 547 EENRGAVIFHGASQDIAPSQSORIASQPSRIQYFGQSLSGQDGLTMDGLVDLAGV 606
QY 606 AQHVLILRSQVLRVKALMERNPREVARNFECNDQVVKGBAGVRVCLHVQKSTRDR 665
Db 607 SKGRVLLLRTPILRVSPVHFTPAEISRVFECQVAPQEQTLSDATVCLLHIESPKTQ 666
QY 666 LREGQIOSVVTYDLADSGRPHSRVFNKTSRQTVLGLTQTCETLKLQLPNCIED 725
Db 667 L--GDURSTVTFDLADHGRSLSTRAIFKETKTALTRVKTGLNKHCHESVSKULLPACVD 724
QY 726 PVSPIVLRNLSLVGTPSLAFGNLRPVLAEDAQLFTALPPPEKNGCNDNI CODDLSITF 785
Db 725 SVTPTILRNLSLVGVPIESLQNLQPLAVDDQTYFTASLPPEKNGADHI CODDLSVVF 784
QY 786 SFMSLDCLVVGPREPNVTVTVRNDGEDSYRTQVTFPFLDLSYRKVSTLQ----- 836
Db 785 GFPDLKTLVVGSDLELNVDTVSNDCEDSYGTTVTLFYPVGLSFRFVABEQVFLRKKEDQ 844
QY 837 --NORSQSWRLACESASSTVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGN 894
Db 845 QWRRQGHSLHLMCD--STPDRSQGLWSTSCSRHVIIFRGSGQMTFLVTFDVSXPKAELGD 902
QY 895 KULLKANVTSENMPRTNKTFOLELPVKVAVVTVVTSHGVSXKYLNFETASE-NTSRVMO 953
Db 903 RULLLRARVGSENNVPCTTKITFOLELPVKVAVVTVVTSIHSHDQTKYLNFTSEKETSVE 962
QY 954 HQYQVSNLQORSPLISLVLPVRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLA 1013
Db 963 HRFQVNNLQORDVPVSINFVWPIELKGEAVV-TVMVSHPNQNPILTQCVRNLRKPTQFDLLT 1021
QY 1014 ELRKAPVNCSTAVCQRIQCDIPFFGIQEEFNATLKNLSFDWYKTSNHLILVSTABI 1073
Db 1022 HMOKSVPLDCSADCLHLRCDIPSLGILDDELFLKGNLSFGWISQTLQKVKLLSEAEI 1081
QY 1074 LFNDVSFTLLPGGAFVRSOTETKVEPEVPNPLPLIVGSSVGGLLLLALITAAALKGLF 1133
Db 1082 TENTSYSQLPQOAEFLRAQTKVLEMYKVHNPVPLIVGSSVGGLLLLAITAILYKGF 1141
QY 1134 FKROYKDMSE 1144
|||||:|
```

```
Db 1142 FKROYKEMLEE 1152
RESULT 6
QYQYE7 PRELIMINARY; PRT; 1161 AA.
ID Q9QYE7
AC Q9QYE7;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RT Gallatin W.M.;
RL "Cloning of rat alpha D, a novel beta 2 integrin.";
DR HMBL; AF011334; AAF21241.1; -.
DR HMBL; F11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int. alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.1%; Score 3282.5; DB 11; Length 1161;
Best Local Similarity 58.1%; Pred. No. 4.7e-231;
Matches 663; Conservative 158; Mismatches 306; Indels 15; Gaps 10;

QY 5 VLLLT--ALTLCHEFNLDTENAMTFQENARFGQSVQVQSGVVGAPQEIIVAAANQSGS 62
Db 6 VILLCGWLASCHGSLNLDVEEPIVREDAAASFGQTVVQFGSRLVVGAPLEAVAVNQTR 65
QY 63 LYCCDYSGSCPEIRLQVPVAVNMSLGLSLAATTSPPQLLACGTVHQTCSNTYVVKGL 122
Db 66 LYDCAPATCMCOPIVLRSPLEAVNMSLGLSVTATNNAQLLACGTAQORACVKNMYKGS 125
QY 123 CFLFGSLNRQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFEVSTVMEQLKK 182
Db 126 CLLLGSLL-QITQAVPASMPECPRQEMDIAFLIDGSGSINORDFAQMDFVKALMGEPAS 184
QY 183 SKTLFSLMOYSBEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFINIT 242
Db 185 TSTLFLMOYSNLKTHFTTFEKNLDPQSLVDVQLQGLTYTATGIRTWHEELPHSK 244
QY 243 NGARKNAPKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKOELNT 302
Db 245 NGRSKAKILLVITDQKYPDPLEYSDVIPAADKAGIIRYAIGVGDAPFOEPTALKELNT 304
QY 303 IASKPPRDHVFQVNFPEALKTIQNLREKI FAIEGTQTCSSSFEHEMSQEGFSAITNS 362
Db 305 IGSAPPQDHVFKVGNFAALRSIQRLQBEKIFAIEGTQSRSSSFQHEMSQEGFSALTSD 364
QY 363 GPLLSTVGSYDWAGGVFLYTSKESKTFINMTRVDSMDNDAYLGAAAIILNRVQSLVIG 422
Db 365 GPVLGAVGFSFGAGFLYPPNTRPTFINMSQENEDMRDYSYLGYSTAVAFWKGVHSLILG 424
QY 423 APRYQHI GLVAMFRONTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHY 482
Db 425 APRHQHTGKVIIFTQESRHRPKSEVRGTQIGSYFGASLCSDVDVDRDGSITDLVIGAPHY 484
```

```
QY 483 YEOTRGQVSVCLPQRCARWQCDVLYGEOGQWGRFGAALTVLGDVNGDKLTVAIG 542
Db 485 YEOTRGQVSVFPVP-QVRGRWOCEATLHGEQHPWGRFVALTVLGDVNGDNLDVAIG 543
QY 543 APGEEDNRGAVLYPHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGGQDLTMDGLVDL 602
Db 544 APGEESRGAVYIFHGASRLIMPSPSRVTGSQLSLRLQVFGQSLSGGQDLTQDGLVDL 603
QY 603 TVGAQGHVLLRSOPVLVRKAIMEFNPREVARNVFECDNQVVGKEAGEVRVCLHVOKST 662
Db 604 AVGAQGHVLLRSPLKLVLSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGS 663
QY 663 RDLREGQISVTVYDALDGRPHSRVAVNETKNSRRTQTVGLGTQTCETIKLQLPNC 722
Db 664 PDLL--GNVQSVRYDALDGRRLISRAIFDETKNCTLTGRKTLGLGDHCEIVKLLPDC 721
QY 723 IEDVSPVLRNLNLSVGTPLSAFNGRLPVLAEADQRLFTALFFPERKNCNDNICDDLS 782
Db 722 VEDAVSPILRLNLSVYDSASP-RNLHPVLAVGSQDHIITASLPFEKNCQELLCEGLG 780
QY 783 ITFSFMSLDCVLVGGPHEFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSR 842
Db 781 ISFNFGQLVVLVGGSPBELTVTVTMNEGEDSYGLTVKFYYPAGLSYRRVTGTQ-QPHQY 839
QY 843 SWRLACESASSTEVSALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANV 902
Db 840 PLRLACEAPAAQED--LRSSSCSINHPIFREGAKTTFMIITFDVSKAFGLDRLLLRKA 897
QY 903 TSENMPRTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFTASNTSR-VNQHOYQVSNL 961
Db 998 SENKNPDNTKTAQLELPVKYTVYLLISROEDSTNHVFNSSHGRRQEAHRYRNNL 957
QY 962 GQSLPSLSVLPVRLNQVWDRPQVTFSENLS--TCHTKERLPHSDFLAELKAPV 1020
Db 958 SPLKLAVRNFWPVLINGVAVMD---VTLSSPAQGVSCVSMQKPPQNPDLTQIQRSS 1014
QY 1021 VNCISAVCORQCDIPFPGIOEENATLKGNLSPDWYIKTSHNLLIVSTAEILFNDVSF 1080
Db 1015 LDCSIADCLHFRCDIPSLDIQDELDLILRGNLSPGWVSQTLQOEKVLVSEAEITFDTSVY 1074
QY 1081 TLLPQCAFVRASQETKVEPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKD 1140
Db 1075 SOLPQCAFRAQVETLLEEVVVEPIFLVAGSSVGGLLLLALITVLYLKLGFKKROYKE 1134
QY 1141 MM 1142
Db 1135 ML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TissUE=Spleen;
RA Ma R. Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF065902; AAD25885.1;
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
```

```
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFDOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWF; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR SEQUENCE 1161 AA; 128240 MW; 868102P7B209E431 CRC64;

Query Match 25.9%; Score 1545; DB 11; Length 1161;
Best Local Similarity 34.0%; Pred. No. 8.5e-104;
Matches 407; Conservative 219; Mismatches 462; Indels 108; Gaps 39;

QY 1 MALRV---LLLTALT---CHGNLDTENAMTFQENA-RGFGQSVVQLQSGRVVVGAPQ 52
Db 1 MSFRIAGPRLLLGLGLFAKAMSYNLDTRPTQSLAQGRHFGYQLVIEDG-VVVGAPG 59
QY 53 EIVAAHQRSGLYOCYSTGSCPEIRLQVPVVEAVNMSLGLSLAAATTSPQLLACGPTVHOT 112
Db 60 E---GDNLTGGLYHCRTSSEFCQPVSLH-GSNHTSKYLGWTLATDAAGKSLACDPGLSRT 115
QY 113 CSNTYVYGLCLFLFGNLRQOPQKPFPEARLQRCQEDSDIAFLIDGSGSIIPHDFRRKKEF 172
Db 116 CDQNTVLSGLCYLFPQSLGPMQLQNRPAYQECMKGVLDLFLFDGSGSLDRKDFEKILEF 175
QY 173 VSTVYMQLKSKTSLFSLMQYSEEFRIHFTFKEP-QNNPNRSLVKPITQLLGRTHATGI 231
Db 176 MKDVMRLKNTSYQFAAVQSTDCRTEFTPLDVYKQNKNDPDLVLLGSVQPMFLTNTPTAI 235
QY 232 RKVVRLEFNITNGARKNAFKILVVIITDGEKFGDPLGYEDVIEADREG-----VIRYV 284
Db 236 NYVVAHVFEESGARPDATKVLIIIDG-----EASDKGNISAAHDITRYI 281
QY 285 IGVGDAPRSKSKSRELNTIASKPPRDHVFQVNNFEAKTIONQLREKIFAIEGTQTGSSS 344
Db 282 IGIGKHPVSQVKQKTLHIIFASEPVEEFVKILDTFEKLDLFTDLQRRYIAIEGTNRQDLT 341
QY 345 SFHEMSQESFAAISNGPLLSSTVGSYDWAGGVF-LYTSKESKSTFNTMTRVSDMNDAY 403
Db 342 SFNMLSSSGISADLSKGHAVGAVGAKWAGGFLDREDLQGGATFVGGPPLTSDVGGY 401
QY 404 LGYAAA-IILNRNVQSLVLGAPRYQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGAS 460
Db 402 LGYTVAMTSSRSRPLAAGAPRYOHVQVLLFOAPEAGGRWNQTKIEGTQIGSYFGE 461
QY 461 LCSVDVDSNGSTDVLIGAPHYIEQTRGGQVSVCLPRGORARWQCDVLYGEOGQWGR 520
Db 462 LCSVDLDQDGEAEILLIGAPLFFQEQGRGVFTY---QRRQSLFEMVSELQGDGYPILGR 518
QY 521 FGAALTVLGDVNGDKLTDAIGAPEEDNRGAVLYPHGTSGSGISPSHSQRIAGSKLSR 580
Db 519 FGAALTALTDINGRLTDAVAVGAPLEE--QGVYIFNGKPG-GUSPQSRQIQAGVPPG 575
QY 581 LOYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQVPLRVKAIMFNPREVARNVFECDN 640
Db 576 IRWFGRSIHGVKDLGGDRLDADVVGPEGRVVLSSRPVVDVVTLSFSPESIPVHEVECS 635
QY 641 DVVYKGEAG-EVRVCLHVOKSTRDLRLREGIOQSVTVYDLDALDGRPHSRVAVNETKNS 699
Db 636 YSAREEQKHGVKLKACFRKPLTPQ--FQGRLLANLSYTLQLDGHRMRSRGLFPDGSHEL 693
QY 700 RRQTVGLGTQTCETIKLQLPNCIEDPSPVILPLNLSLV---CTPLSAFQN-LRPVLAE 755
Db 694 SGNTSITP-DKSCLDLPHFPIICIDLISPINVSFLSLEEEGTPRQKRAMQPIILRP 752
QY 756 DAQRLFTALPPEKNCNDNICDDLSITFSFMSLDCVLVGGP-----REFNVTVTVRN 809
Db 753 SIHTV-TKEIFEKNCGEDKCEANLITSSPARS-----GPLRLMSSASLAVETWLSN 804
QY 810 DGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWRLACESASSTEVSAL-KSTSCSIN 868
Db 810 DGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWRLACESASSTEVSAL-KSTSCSIN 868
```

Db	*805	SGEDAYVVRDLDDPPRGLSFRKVEMLQ----	PHSRMPVSCBEL-----TEGSSLLTKLKCNVS	859
Qy	869	HPFIPENSEVTNITFOVDSKASIGNKLLKANVTSEN-NMPRNTKTEFOLELPVKYAVY	927	
Db	860	SPFKAQGEVSLQVMFPTLLNSSWEDFVELNGTVHCENENSSLOEDNSAATHIPVLVPVN	919	
Qy	928	MVVTSHGVSTKYLNFNTASENSTRVMOHQYOY----	SNLQGRSLPISLVLFPVPLNQTVTW	984
Db	920	ILTKEQENSTLYISFTPKGPKTOQVHQVYQVRQPSAYDHNMP-TLEALVGVP-----W	972	
Qy	985	DRPQ--VTFPSNYS-----TCHTKB-RLPSSHSDFLAELRKAPVYVNCISAVCQRIQCDIPF	1037	
Db	973	PHSBDPITYTWSVQTDPLVTCHESDLKRPSE---AEQPCLPGV-----QFRCPIVF	1021	
Qy	1038	FGIOEEFNATLKGNLSFDWIKTSHNHLILVSTAEILLFNDISVFTLLPGQCAFVRSQTETK	1097	
Db	1022	---RRELIQVTGVTELSEKIKAS-STLSICSSLVSUSFNSSKUFHLYGSKA-SEAQVLVK	1076	
Qy	1098	VEPEVPNPPLIVGVSSVGGLLLIALITALYKLGFFKROYKDMW-SEGPPGGAEP	1152	
Db	1077	VDLTHEKEMHLHVYVLSGIGGLVLLFLFLAYLYKVGFFKRLKEKMEADGQVNGSP	1132	

RESULT 8

Q9R200	PRELIMINARY;	PRT; 1160 AA.
ID	Q9R200	
AC	Q9R200;	
DT	01-MAY-2000 (Tremblrel. 13)	Created
DT	01-MAY-2000 (Tremblrel. 13)	Last sequence update)
DT	01-JUN-2002 (Tremblrel. 21)	Last annotation update)
DE	Integrin alpha L.	
GN	ITGAL.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa;	Cnidata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia;	Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN=CS7BL/6J; TISSUE=Spleen;	
RA	Ma R.Z., Teuscher C.;	
RL	Submitted (MAY-1998) to the	EMBL/GenBank/DBJ databases.
DR	EMBL; AF065901; AAD25884.1;	
DR	HSSP; P20701; ILFA.	
DR	MCD; MG1.96606; Itgal.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01839; PG-GAP; 5.	
DR	Pfam; PF00357; integrin_A; 1.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 5.	
DR	SMART; SM00327; VWA; 1.	
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR	PROSITE; PS50234; VWF; 1.	
DR	SEQUENCE 1160 AA; 128127 MW;	A33C531B139F1FAD CRC64;
SQ	SEQUENCE	

Qy	173	VSTMOLKSKTKTLFSLMQVSEBFRHFTPFKEF-QNNPNRSLVKPITQLLQGRTHATGI	231
Db	176	MKDWMRKLNTSYQFAAVQFSTDCRTEFTFLDYVVKQNKNDPVLGSGVQPMFLTLNTFRAI	235
Qy	232	RKVRELFNITTCARKNAFKILVITDGEKFGDPLGVEDVIPEADREG-----VIRYV	284
Db	236	NYVVAHVFEESGARPDATKVLVIITDG-----BASQKGNISAAHDIITRYI	281
Qy	285	IGVGDAFRSEKSRQELNTIASKPPRDRHVFOVNNFEALKTIQNLORKEIFAIEGTOGSSS	344
Db	282	IGIGKHFVSVQKQKTLHIIFASEPVEBFVKILDTFEKLIKOLFTDLQRRYIAIEGTRQDLT	341
Qy	345	SPHEMSQBGFSAAITNSGPELLSTVGSYDWAGVFLYTSKEKSTFINNTRVDSNDNDAY	403
Db	342	SFMNELSSSGISADLSKGHAVGAVGAKWAGGFLDLREDLOGATFVGQEPILTSVRGGY	401
Qy	404	LGVAAA-IILNRNVQSILVLAGAPRYQHIGLVAMFR--QNTGMWESNANVKGTOIGAYFGAS	460
Db	402	LGVTVAMTSTRSRPLLAAGAPRYQHVGQVLLFQAPEAGGRMNQTKIEGTIGSYFGGE	461
Qy	461	LCSVDVDSNGSTDVLITIGAPHYYEQTRGGOVSVCPPLRGQARWQCDAVLYGQGOPMWR	520
Db	462	LCSVDLDODGEAEALLIGAPLFFGEQRGRVFY---QRRQSLFEMVSELQGDGPLYLGR	518
Qy	521	FGAALTVLGVNDGDKULTVAIGAPEGBEDNKGAVYLFHGTSGSGISPSHSQRIAGSKLSPR	580
Db	519	FGAAITALTIDINGDRULTDVAVGAPLEE--QGAVYIFNGKPG-GLSPQSPQRIQGAQVFPG	575
Qy	581	LQYFGQSLSCGDLTWDGLVDLTVGAQGHVLLRLSQPVLVKAIMFENPREVARNVFECH	640
Db	576	IRWFGRSIGHVKDGLGDRLANVVGPEGRVVLLSRPVDVVTLELFSPEEIPVHEVECS	635
Qy	641	DQVVKGEAG-EVYRVCLHVQKSTRDLRECOIQSVVTVYDIALDSGRPHSRAVFNETKNST	699
Db	636	YSAREEQKGVKLKACERIKPLTPQ--FQORLLANUSYTLQLOGHRMRSKCLPFGDSHEL	693
Qy	700	RROTQVLGLTOTCETUKLQLPNCIEDPVSPIVLRNPSLV--GTPUSAFGN-LRPVLAE	755
Db	694	SGNTSITP-DKSCLDHFHFHFPICIQULISPIVNSLNFSLLEEBGTPRDQKGRAMQPILRP	752
Qy	756	DAQRLFTALPFPKNCNDNIQQDLSITFSFMSLCLLVGGP-----REFNVTVTVRN	809
Db	753	SIHTV-TKEIPFKNCGEDKKCEANULTSSPARK-----GURLMSSASLAVEMWTLN	804
Qy	810	DGEDSVRQTVTFPPDLDLSYKRVSTLQNRQSRSWRLACASASTEVSAGAL-KSTCSIN	868
Db	805	SGEDAYVWRDLDFPCRLSFRKVMLO--PHSRMPVSCBEL--TEGSSLLTKLTKCNVS	859
Qy	869	HPFIPENSEVTFNITFDVDSKASLGKLLKANVTSN-NMPTNKTTEFOLELPVKYAVY	927
Db	860	SPIFKAGQSVSLQVMFNTLLNSWEDFVELNGTVHCENENSSLOEDNSAATHIPVLVFN	919
Qy	928	MVYTSCHVSTKYLNFASENTSRVMQHOYQVSNLQORSPLISLVFLVPVRLNQTIVWDRP	987
Db	920	ILTKQENSNLYISFTPGPKTQQQVHYQV----RIQPSAYDHNMT-LEALVGVPRP	973
Qy	988	Q-----VTFSENLS---TCHTKE-RUPSHSDFLAELRKAPVNCSTAVCORICQDIPFF	1038
Db	974	HSEDLIYTVSWQTDPLVTCHSEDLKRP-----SEAPCLPGV--QFRCPIVF-1020	
Qy	1039	GIQEEFNATLKGNSFDWYIKTSHNHLLIYSTABIPLFNDSPVFTLLPQOGAFVRSQTETKV	1098
Db	1021	--RWEILLIQTGTVELSKEIKAS-STLSCSSLSVSPNSSKHFLHYGSKA-SEAQVLVKV	1076
Qy	1099	BPTEVPNPLPIVGVSSVGGILLALITAAALYKLGFFKRYKQKDM-SSGGPPGABP	1152
Db	1077	DLTHEKMLHVYVLSGIGGLVLLFLIFALYKVGFFKRNLEKMEADGGVNGSP	1131

RESULT 9

Q98TF1

ID Q98TF1

AC Q98TF1;

PRELIMINARY;

PRT; 1196 AA.

RESULT 9
Q98TF1
ID Q98'
AC Q98'


```
QY 175 TVMEQLKSKTFLSMOYSBEFRTHFTKEFQNNPNRSLVKPITQLLGRTHATGIRKV 234
Db 108 -----ILPAVQFSYKTEFDPSDVYKRPDALLXVHKWMLLTNTFGALNYV 157
QY 235 VRELNTNGARKNAFKILVVITDGEKPGDPLGYEDVIPADREGVIRYVIGVDAPRSE 294
Db 158 ATEVFREELGARPDAKVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIKHFKQTK 210
QY 295 KSRELNTIASKPRDHVQVNNFEALKTIQNLREKIFALEGTQGTSSSFEHMSQEG 354
Db 211 ESQETLHKFKPASEFVKILDTFEKLKDLTFELQKKIYVIEGTSKQDLTSFNNELSSSG 270
QY 355 FSAAITSGPLLSVGVSDWAGGVF-LVTSKEKSTFINMTRVDSMDMDAYLGAAA-IIL 412
Db 271 ISADLSRGHVVAGAKDWAGGFLDKADLDDTFIGNEPLEPVEVRAGLYGVTVWLPS 330
QY 413 RNRVQSLVLGAPRYOHLGLVAMFR--QNTGMESNANVKGTOIGAYFASLCSVDVDSNG 470
Db 331 ROKTSLLASGAPRYOHRGVRLLFQBPQGGHWSQVQTHGTQIGSYFGGELCGVDVDQD 390
QY 471 STDVLIGAPHYYEQTRGGQVSVCLPRGQARWQCDV--LYGEOGQPHRGCAALTVL 528
Db 391 ETELLIGAPLYFGEOGRGVFIY-----QRRQLGPEVSELOQDGPYGLRFGAITAL 445
QY 529 GDVNGDKLTDVAIGAPGEDNRGAVYLPFGHTSGSGISPSHSORTAGSKLSPRLQYQOSL 588
Db 446 TDINGDGLVDVAVAGAPLEE--QGAVYIPNGRHG--GLSPQPSQRLEGTVLSGIGWFGRSI 502
QY 589 SGGQDLTMDGLVDLTVAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQV--KGK 647
Db 503 HGVKLEGDGLADVAVAGAESQMVLLSRPVDMVTLMSFSPAEPVHEVECSYSTSNKMK 562
QY 648 EAGEVRVCLHVQKSTRDLREGQTSQVVTYDIALDSGRPHSRVAVNETKSTRQTOVLG 707
Db 563 EGVNITICFQT-KSLIPQF-QGRLVANITYTLQDGHRTRRGLFPGGRHELNRNIAYT- 619
QY 708 LTQTCETLKLQPLNCIEBPVPIVLRNLFSL---VGTPLSAFGN-----LRPVLAEQAQ 758
Db 620 TSMCTDPSFHPVCVQDLISPIVNSLNFSLWELEGTRDQAGKDIPPLRPSLHSETW 679
QY 759 RLFTALPFPERKNCNDNF CODLSITFSFMSLDCLVGGPREFNVTVVRNDESDSYRTQ 818
Db 680 EI-----PFKXCGEDKCEANLRVSPARSRALRLTAFASLSVELSLSNLEBDYVWQ 734
QY 819 VTFPPDLSTRKUSTLQNRQSRWRLACES--ASSTEVSGALKSTSCSINHIPPENS 876
Db 735 LDLHFPPLGLSPKVEML--KPHSQIPVSCBELPEESRLLSRAL---SCNVSSPIFKAGH 788
QY 877 EVTFNITFDVDSKASLGNKLLKANVTSENN---MPRTNKTFOLELPVAVYVMVVT 932
Db 789 SVALQMMNTLVNSWGSVELHANVTCWNEDSDLEDNSATTI---IPILYPINILIOD 845
QY 933 HGVSTKYLNFASENTSVMQHQYQV---SNLQGRSLP-ISLVLVPLVRLNQTWIMDRPQ 988
Db 846 QEDSTLYVSFTPKPKIHOVKHMYQVRIQPSIHDHNIPTLEAVGVPPQPPSEGPITHQWS 905
QY 989 VTFSENLSSTCHTK--ERLPHSD--FLAELRKAPVNCSTAVCORIQDIPFGIQEFP 1044
Db 906 VQMEPPV--PCHYEDLERLPDAAEPCLPALFCFV-----FROEI 945
QY 1045 NATLKNLSFDMYIKTSNHLHIVSTAELFNDSVFTLLPGQAFVRVSRQTEKVEPPEVP 1104
Db 946 LVQVIGTVELVGEIEAS-SMFSLCSLSISFNSSKGFHLYGSNASL-AQVWVKVDVVEK 1003
QY 1105 NPLPLIVSSVGGLLLLLALITAAALYKLGFFKRYQKDMMSSEG-GPPGAEP 1152
Db 1004 QMLYLYVLSGIGLLLLLIIFVLVYKGVFFKENLKEKMEAGRGVNGIP 1052
RESULT 12
Q8HZVO
ID PRELIMINARY; PRT; 927 AA.
```

```
AC Q8HZVO;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1];
RP SEQUENCE FROM N.A.
RA Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Sequence of the alpha subunit of bovine lymphocyte function-
RT associated antigen 1.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF40778; AN63636.1; -.
FT NON_TER 1
FT NON_TER 927
SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match 21.4%; Score 1277; DB 6; Length 927;
Best Local Similarity 34.8%; Pred. No. 2.5e-84;
Matches 340; Conservative 167; Mismatches 385; Indels 86; Gaps 28;

QY 190 MOYSEBFRIHFTKSFQNNPNRSLVKPITQLLGRTHATGIRKVRELFNITNGARKNA 249
Db 4 VQSTVYFRTFTFLDIYIRKQDPDALLAGVHKWMLLTNTFGAIVYVAKEVFRPDLGARPD 63
QY 250 FXILVITDGEKPGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQELNTIASKPPR 309
Db 64 TKVLIITIDGK--PPTNTTLMRPKTS-----RSLIGIKNFYTKESQEAALHOFASKPVE 116
QY 310 DHVPQVNNFEALKTIQNLREKIFALEGTQGTSSSFEHMSQEGFSAATITSGNPLLSV 369
Db 117 EFVKILDTFEKLKDLTFELQKKIYVIEGTSKQDLTSFNNELSSGISADLSSEGHVGVAV 176
QY 370 GSYDWDAGGVF-LVTSKEKSTFINMTRVDSMDMDAYLGAAA-IILNRVQSLVLGAPRYQ 427
Db 177 GAKDWAGGFLDLKADLKSSTFVGNELPTVESRAGLYGVTVRLPSRGTMSLLATGAPKYQ 236
QY 428 HIGLVAMFR--NTGMESNANVKGTOIGAYFASLCSVDVDSNGSTDVLIGAPHYYEQ 485
Db 237 HVGRVLLFQPKRGGPMQIQEIDGIGSYFGGELCGVDVDRDGETELLIIAAPLYGE 296
QY 486 TRGQSVSVCLPRGQARWQCDVLYGEOGQPHRGCAALTVLGDVNGDKLTDVAIGAPG 545
Db 297 QRGGRVFIY--OKIQLEFQMVSELOGETGYPLGRFGAAIATLTDINGDELTDVAVGAPL 353
QY 546 EEDNRGAVYLFHGTSGSGISPSHSORTAGSKLSPRLQYFQCSLGGQDLTMDGLVDLTVG 605
Db 354 EE--QGAVYIFNGQOG--GLSPRPSQRLEGTMFSGIQMFGRSIGHVKDGLGGDLADVAVG 410
QY 606 AQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVVGKKEAG-EVRVCLHVQK--ST 662
Db 411 AEGQVIVLSSRPVVDIIITSVSFSPAEPVHEVECSYSTSNQKKEGVNLTVCFOVKSLIST 470
QY 663 RDLRREGQIOSVVTYDIALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQPLNC 722
Db 471 -----FQGHVLANITYTLQDGHRTSRGLPFGGKHKLIGNTAVTPV-KSEVFVHFHPIC 525
QY 723 IEDPVSPIVLRNLFSL---VGTPLS--AFGNLRVPLAEADAQRLFTALPFPERKNCNDNIC 777
Db 526 IQDLISPIVNSLSYSLWELEGTRDPRALDRDIPPLKPSHLETKETIPPERKNCGEDKNC 585
QY 778 QDDLISITFSFMSLDCLVGGPREFNVTVVRNDEGDSYRTQVTFPFDLSYRKVSTLON 837
Db 586 EADLKLAFSDMRSKILRLTPSASVLRLTARNTAEDAYVQVTLSPFGGLSFRKVEIL-- 643
QY 838 QRSQSRWLACESASSTEVSGALKSTSCSINHIPPENSEVTFNITFDVDSKASLGNKLL 897
Db 644 -KPHSHVPVGCCELPEEBAVHVS-RALSCNVSSPIFGEDSMVDIQVMNTLQKSGWGFIE 701
```


Search completed: November 25, 2003, 14:20:22
Job time : 39.5727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.8086 seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MALRVLLTALTLCHEFNLD.....FKRQYKDMSEGGPGGAEPPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	AAW65090	Human Beta-integrin
2	5956	100.0	1153	ABU07406	Human CD11b protei
3	5956	100.0	1153	ABG61469	Human Beta2 integr
4	5956	100.0	1153	AAU80252	Human integrin 1 a
5	5956	100.0	1153	AAO14428	Integrin Mac-1 alp
6	5946	99.8	1153	AAAR04136	Alpha subunit of M
7	3494	58.7	1163	AAAR07120	p150.95 alpha subu
8	3480	58.4	1163	AAW65091	Human Beta-integrin
9	3480	58.4	1163	ABU07361	Human CD11c protei

10	3480	58.4	1163	23	ABG61470	Human Beta2 integr
11	3478	58.4	1163	24	ABU07406	Protein different
12	3455	58.0	1161	16	AAW78166	Human beta-2 integ
13	3455	58.0	1161	18	AAW23049	Human beta 2 integ
14	3455	58.0	1161	19	AAW72825	Human alpha-d. Ho
15	3455	58.0	1161	19	AAW65089	Human Beta-integrin
16	3455	58.0	1161	19	AAW57491	Human Beta2 integr
17	3455	58.0	1161	20	AAW73342	Human alpha d clon
18	3455	58.0	1161	21	AAW07359	Human alpha d clon
19	3455	58.0	1161	21	ABG61468	Human Beta2-integr
20	3439.5	57.7	1161	18	AAW23064	Human beta 2 integ
21	3439.5	57.7	1161	19	AAW2837	Human alpha-d deri
22	3439.5	57.7	1161	19	AAW65106	Human Beta-integrin
23	3439.5	57.7	1161	20	AAW73343	Human alpha d prote
24	3439.5	57.7	1161	21	AAW07376	Human alpha d prote
25	3439.5	57.7	1161	21	ABG61485	Human Beta2-integr
26	3272.5	54.9	1161	16	AAW78169	Rat alpha-d subuni
27	3270.5	54.9	1161	18	AAW23062	Rat beta 2 integri
28	3270.5	54.9	1161	19	AAW72824	Rat alpha-d #1. R
29	3270.5	54.9	1161	19	AAW60004	Rat alpha d polype
30	3270.5	54.9	1161	21	AAW07374	Rat alpha d protei
31	3270.5	54.9	1161	21	ABG61483	Rat Beta2-integrin
32	3264	54.8	1161	18	AAW23061	Mouse beta 2 integ
33	3264	54.8	1161	19	AAW72836	Mouse alpha-d #2.
34	3264	54.8	1161	19	AAW65103	Mouse beta-integrin
35	3264	54.8	1161	19	AAW60003	Mouse alpha d poly
36	3264	54.8	1161	20	AAW73347	Mouse alpha d prote
37	3264	54.8	1161	21	AAW07373	Mouse alpha d prote
38	3264	54.8	1161	23	ABG61482	Mouse Beta2 integr
39	3263.5	54.8	1161	19	AAW65104	Rat beta-integrin
40	3263.5	54.8	1161	20	AAW73345	Rat alpha d subu
41	3260	54.7	1161	16	AAW78168	Mouse alpha-d subu
42	3251.5	54.6	1151	18	AAW23059	Rat beta 2 integri
43	3251.5	54.6	1151	19	AAW72834	Rat alpha-d #2. R
44	3251.5	54.6	1151	19	AAW65101	Rat beta-integrin
45	3251.5	54.6	1151	19	AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

DT 28-SEP-1998 (first entry)

DE Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
XX lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX rheumatoid arthritis.

OS Homo sapiens.

PN US5728533-A.

XX 17-MAR-1998.

PF 07-JUN-1995; 95US-0485618.

XX 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 05-AUG-1994; 94US-0286889.

PR 21-DEC-1994; 94US-0362652.

XX (ICOS-) ICOS CORP.

XX Gallatin WM, Van DER VIEREN M;

DR XX WPI; 1998-206565/18.

PT Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-1 diabetes

XX

PS Example 5; Fig 1A-D; 106pp; English.

XX

CC This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX

SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 19; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTCHGPNLDNTENAMTFQENARGFGQSVVQLQGRVVVGAPQEIIVAAANQR 60

DB 1 MALRVLLLTALTCHGPNLDNTENAMTFQENARGFGQSVVQLQGRVVVGAPQEIIVAAANQR 60

QY 61 GSLVQCDYSTGSCPEIRLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120

DB 61 GSLVQCDYSTGSCPEIRLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120

QY 121 GLCFUFGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQL 180

DB 121 GLCFUFGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQL 180

QY 181 KKSRTLFLSMQYSEBFRTHFTKFEQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240

DB 181 KKSRTLFLSMQYSEBFRTHFTKFEQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240

QY 241 ITNGARKNAFVLVITGEGFDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQEL 300

DB 241 ITNGARKNAFVLVITGEGFDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQEL 300

QY 301 NTIASKPPRDHVFQNNFEALKTIONQREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360

DB 301 NTIASKPPRDHVFQNNFEALKTIONQREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360

QY 361 SNGPLLSTVGSYDAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRNVQSLV 420

DB 361 SNGPLLSTVGSYDAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRNVQSLV 420

QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480

DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480

QY 481 HYYEOTRGQGVSVCPPLRGQARWOCDAVLGEQGPWGRFGAALTVLGVNGDKLTDVA 540

DB 481 HYYEOTRGQGVSVCPPLRGQARWOCDAVLGEQGPWGRFGAALTVLGVNGDKLTDVA 540

QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFQGSLSGGQDLTMDGLV 600

DB 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFQGSLSGGQDLTMDGLV 600

QY 601 DLTVCAQGHVLLRSPQVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOK 660

DB 601 DLTVCAQGHVLLRSPQVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOK 660

QY 661 STRDLRREGIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVTLGLTQTCETLKLQLP 720

DB 661 STRDLRREGIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVTLGLTQTCETLKLQLP 720

QY 721 NCIEDPVSPVILRLNFSLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKXKCGNDNICODD 780

DB 721 NCIEDPVSPVILRLNFSLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKXKCGNDNICODD 780

QY 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNR 840

DB 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNR 840

QY 841 QRSWLACSSASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

DB 841 QRSWLACSSASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

QY 901 NVTSNNMPTNKTETQLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOVSN 960

DB 901 NVTSNNMPTNKTETQLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOVSN 960

QY 961 LGORSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020

DB 961 LGORSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020

QY 1021 VNCISIAVCQRIQCDIPFFGIOEEFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDVSF 1080

DB 1021 VNCISIAVCQRIQCDIPFFGIOEEFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDVSF 1080

QY 1081 TLLPGQGAFAVRSQETKVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKD 1140

DB 1081 TLLPGQGAFAVRSQETKVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKD 1140

QY 1141 MMSEGGPPGAEPO 1153

DB 1141 MMSEGGPPGAEPO 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX

AC AAB07360;

XX

DT 17-JAN-2001 (first entry)

XX

DE Human CD11b protein sequence.

XX

KW Human; macrophage infiltration inhibition; alpha_d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX

OS Homo sapiens.

PN WO200029446-A1.

XX

PD 25-MAY-2000.

XX

PF 16-NOV-1999; 99WO-US27139.

XX

PR 16-NOV-1998; 98US-0193043.

PR 08-JUL-1999; 99US-0350259.

XX

XX (ICOS-) ICOS CORP.

PA Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

XX

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites

XX

PS Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha_d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha_d. The Alpha_d gene and protein may be useful in therapy for diseases linked to aberrant alpha_d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

XX Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 21; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAURVLLLTALTLCGFNDLTENAMTFQENARFGOSVVOLOGSRVVVVGAPQEIIVAAVOR 60
DB 1 MAURVLLLTALTLCGFNDLTENAMTFQENARFGOSVVOLOGSRVVVVGAPQEIIVAAVOR 60
QY 61 GSIYQCDYSTGSCERILQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVK 120
DB 61 GSIYQCDYSTGSCERILQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVK 120
QY 121 GLCFLFGSNLRQOPKPPALRCQPOEDSDIAFLIDGSGSIIPHDPRRMKEFVSTVMEQL 180
DB 121 GLCFLFGSNLRQOPKPPALRCQPOEDSDIAFLIDGSGSIIPHDPRRMKEFVSTVMEQL 180
QY 181 KSKTLFSLMOYSEERIHFTFKFONNPNRSLVKPITOLLGRHTATGIRKVVRELFN 240
DB 181 KSKTLFSLMOYSEERIHFTFKFONNPNRSLVKPITOLLGRHTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAIT 360
DB 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAIT 360
QY 361 SNGPLLSTVGSYDWAGVFLYTSKESKSTFINMTRVDSMDNDAYLVYAAAILLRNVQSLV 420
DB 361 SNGPLLSTVGSYDWAGVFLYTSKESKSTFINMTRVDSMDNDAYLVYAAAILLRNVQSLV 420
QY 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEQTGRGGQSVCPPLPRGQARWQCDVAVLYGSGQPGWRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEQTGRGGQSVCPPLPRGQARWQCDVAVLYGSGQPGWRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTWDGLV 600
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTWDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDALDGRPHRAVFNFTKNSTRROTQVLGLTQTCTETLKQLP 720

DB 661 STRDLREGQIQSVVYTDALDGRPHRAVFNFTKNSTRROTQVLGLTQTCTETLKQLP 720
QY 721 NCIEDPVPSPVILRLNFSLVGTPLSAFGLNLRPVLAEQAORLEFALFFFKNGCNDNICODD 780
DB 721 NCIEDPVPSPVILRLNFSLVGTPLSAFGLNLRPVLAEQAORLEFALFFFKNGCNDNICODD 780
QY 781 LSITSPFMSLCLVVGSGPREFNVTVVRNDGEDSVRTQVTFPPDLSDYRKVSTLQONRS 840
DB 781 LSITSPFMSLCLVVGSGPREFNVTVVRNDGEDSVRTQVTFPPDLSDYRKVSTLQONRS 840
QY 841 QRSWELACESASSTEVSGALKSTSCSINHPIFFPENSEVTFNITPDVDSKASLGNKLLKA 900
DB 841 QRSWELACESASSTEVSGALKSTSCSINHPIFFPENSEVTFNITPDVDSKASLGNKLLKA 900
QY 901 NYTSENMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQHOYQVSN 960
DB 901 NYTSENMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQHOYQVSN 960
QY 961 LQQRSLPISLVPLVPLVRLNQTVIWDROPVTSSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LQQRSLPISLVPLVPLVRLNQTVIWDROPVTSSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
DB 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
QY 1081 TLLPQGGAFVRSQSTETKVEPFPVNPPLPIVGVSSVGGILLALLAIIAALYKLGFFFKQYKD 1140
DB 1081 TLLPQGGAFVRSQSTETKVEPFPVNPPLPIVGVSSVGGILLALLAIIAALYKLGFFFKQYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153
RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX AC ABG61469;
XX DT 27-AUG-2002 (first entry)
XX DE Human Beta2 integrin alphaCD11b subunit.
XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX PN WO200230980-A2.
XX PD 18-APR-2002.
XX PF 15-OCT-2001; 2001WO-US32059.
XX PR 13-OCT-2000; 2000US-0688307.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX WPI; 2002-463260/49.
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,

or limiting autonomic and sensory dysfunction following spinal cord injury -

Example 5; Page 191-194; 270pp; English.

The invention relates to promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury by administering an anti-alpha d (Beta2) integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha d cDNAs and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia. The present sequence is a Beta2 integrin alpha subunit sequence included for comparison with the Beta2 integrin alpha d protein sequences.

Query Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTCHGFLNLTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQR 60
DB 1 MALRVLLLTALTCHGFLNLTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANQR 60

QY 61 GSLYOCYDSTGSCPIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVK 120
DB 61 GSLYOCYDSTGSCPIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVK 120

QY 121 GLCFLFGLNLRQPPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFEFVSTVMEQL 180
DB 121 GLCFLFGLNLRQPPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFEFVSTVMEQL 180

QY 181 KSKTFLSLMQYSEBFRIHFTKPEFQNNPNRSLVKPTQLLGRTHHTATGIRKVVRELFN 240
DB 181 KSKTFLSLMQYSEBFRIHFTKPEFQNNPNRSLVKPTQLLGRTHHTATGIRKVVRELFN 240

QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDREGVIRYVIGVGDAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDREGVIRYVIGVGDAFRSEKSRQEL 300

QY 301 NTIASKPRDRHVQVNNFEALXTIQNLRKIFAIEGTQTGSSSFEHEMSQEGFSAAIT 360
DB 301 NTIASKPRDRHVQVNNFEALXTIQNLRKIFAIEGTQTGSSSFEHEMSQEGFSAAIT 360

QY 361 SNGPLLTSTVGSVDWAGGVPFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILNRVQSLV 420
DB 361 SNGPLLTSTVGSVDWAGGVPFLYTSKESKSTFINNTRVDSMDNDAYLGYAAAILNRVQSLV 420

QY 421 LGAPRYQHITGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVIGAP 480
DB 421 LGAPRYQHITGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVIGAP 480

QY 481 HYYEOTRGQGVQVCPPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQGVQVCPPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540

QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600

QY 601 DLTVGAGQHVLRLRSQPVLRVKAIMEFNPREVARNVFECDNQVVKGKEAGEVRVCLHVQK 660
DB 601 DLTVGAGQHVLRLRSQPVLRVKAIMEFNPREVARNVFECDNQVVKGKEAGEVRVCLHVQK 660

QY 661 STRDLRREGIQSVVWYDIALDSGRPHSRAVFNETKNSRRQTVGLGLTQTCETLKLQLP 720
DB 661 STRDLRREGIQSVVWYDIALDSGRPHSRAVFNETKNSRRQTVGLGLTQTCETLKLQLP 720

QY 721 NCIEDPVSPIVLRLNFSLVGTPLSAFQNLRLPVLAEADAQRLFTALFFPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRLNFSLVGTPLSAFQNLRLPVLAEADAQRLFTALFFPEKNCNDNICQDD 780

QY 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQVRS 840
DB 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQVRS 840

QY 841 QRSWRLACESASSTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWRLACESASSTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900

QY 901 NVTSENNMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHQVSN 960
DB 901 NVTSENNMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHQVSN 960

QY 961 LGQRSPLISLVFLVPVRLNQTIVDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGQRSPLISLVFLVPVRLNQTIVDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020

QY 1021 VNCIAVCQRIQCDIPFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSF 1080
DB 1021 VNCIAVCQRIQCDIPFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSF 1080

QY 1081 TLLPGQAFVRSOTETKVEPFEVPNPLVIGSVGGLLLLALITAAALYKLGFFKROYKD 1140
DB 1081 TLLPGQAFVRSOTETKVEPFEVPNPLVIGSVGGLLLLALITAAALYKLGFFKROYKD 1140

QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153

RESULT 4
AAU80252
ID AAU80252 standard; Protein; 1153 AA.

XX AC AAU80252;
XX DT 15-JUL-2002 (first entry)
XX Human integrin 1 alpha-M subunit protein.

DE Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjogren's syndrome; rheumatoid arthritis.
XX Homo sapiens.

XX Key Location/Qualifiers
FH Misc-difference 499..500
FT /note= "Encoded by GGG CAG AGG"
XX WO200218593-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27227.
XX 01-SEP-2000; 2000US-229700P.
XX (BLOO-) CENT BLOOD RES INC.
XX Springer TA, Shimoaka M, Lu C;

XX WPI; 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation -
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like
CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilized in a desired conformation. The polypeptide of
CC the invention may have antiinflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis; human immunodeficiency virus (HIV); myocardial
CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.
XX
SQ Sequence 1153 AA;
Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLTLTALTLCHGFNLDENAMTFFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANQR 60
DB 1 MALRVLTLTALTLCHGFNLDENAMTFFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANQR 60
QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLFGSLNROOPKFPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQL 180
DB 121 GLCFLFGSLNROOPKFPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQL 180
QY 181 KKSRTLFLSMQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSRTLFLSMQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVQVNNFALAKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
DB 301 NTIASKPPRDHVQVNNFALAKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
QY 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
DB 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
QY 421 LGAPRYQHITGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHITGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEQTRGGQSVCPILPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDA 540
DB 481 HYYEQTRGGQSVCPILPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDA 540

QY 541 IGAFGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDITMDGLV 600
DB 541 IGAFGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDITMDGLV 600
QY 601 DLTGAQGHVLLLSQPVLRVKAIMEFNPREVARNVFECDNVVKGKAGEVRVCLHVQK 660
DB 601 DLTGAQGHVLLLSQPVLRVKAIMEFNPREVARNVFECDNVVKGKAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRVLAEDAQRFTALFPFEKNCNGNDICODD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRVLAEDAQRFTALFPFEKNCNGNDICODD 780
QY 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDCGDSYRQVTFEFPDLDSYRKVSTLQORS 840
DB 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDCGDSYRQVTFEFPDLDSYRKVSTLQORS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFOLELPVKYAVVMVTSHGVSTKYLNFTASENTRVMOHQYQVSN 960
DB 901 NVTSENMPRTNKTEFOLELPVKYAVVMVTSHGVSTKYLNFTASENTRVMOHQYQVSN 960
QY 961 LGQRSLSPLSLVPLVPRNQTVIWDROPQVTESENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGQRSLSPLSLVPLVPRNQTVIWDROPQVTESENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCISIAVCQRIQCIPPFQIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
DB 1021 VNCISIAVCQRIQCIPPFQIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
QY 1081 TLLPQQAQFVRSQVSTETKVEPEVPNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKD 1140
DB 1081 TLLPQQAQFVRSQVSTETKVEPEVPNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKD 1140
QY 1141 MMSGGPGGAPQ 1153
DB 1141 MMSGGPGGAPQ 1153
RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX
AC AAO14428;
XX
DT 03-MAY-2002 (first entry)
XX
DE Integrin Mac-1 alpha subunit.
XX
KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX
OS Unidentified.
XX
XX WO200204521-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US21805.
XX
PR 07-JUL-2000; 2000US-216600P.
XX

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (BLOO-) CENT BLOOD RES.
XX Springer T;
XX WPI; 2002-148167/19.
XX
XX New integrin I domain protein having alteration in at least 2
XX noncontiguous regions and exits in an open conformation, useful for
XX treating, preventing or suppressing inflammatory or immunological
XX disorders
XX
XX Example 1; Fig 1F; 90pp; English.
XX
XX The invention comprises structurally biased variant integrin inserted (I)
XX domain proteins, wherein the alterations to the protein occur in at least
XX two noncontinuous regions. Specifically the variant integrin I domain
XX proteins are structurally biased to exist in the open conformation,
XX thereby altering the binding ability of the protein. The invention also
XX comprises nucleic acids encoding the variant integrin I domain proteins.
XX The integrin I domain proteins and nucleic acids are useful for treating,
XX preventing or suppressing integrin related inflammatory and immunological
XX disorders (e.g. rheumatoid arthritis). The variant integrin I domain
XX proteins and nucleic acids can also be used for treating: ischaemia/
XX reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
XX infection; and cancer. The variant integrin I domain nucleic acids and
XX proteins may be used in gene therapy, as vaccines and to screen for
XX bioactive agents. The present amino acid sequence represents the Mac-1
XX alpha subunit of integrin.
XX
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQR 60
DB 1 MALRVLLLTALTCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQR 60

QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVK 120

QY 121 GLCLFPGNLRRQPOKFPPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKFEFVTVMEQL 180
DB 121 GLCLFPGNLRRQPOKFPPEARLGCPOEDSDIAFLIDGSGSIIPHDPRMKFEFVTVMEQL 180

QY 181 KSKTTLFSLMOYSEBFRTHFTKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KSKTTLFSLMOYSEBFRTHFTKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240

QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300

QY 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
DB 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360

QY 361 SNGPLLLSTVGSYDWAAGVFLYTSKESKTFINMTRVDSMDNDAYLGAAAIILNRNVQSLV 420
DB 361 SNGPLLLSTVGSYDWAAGVFLYTSKESKTFINMTRVDSMDNDAYLGAAAIILNRNVQSLV 420

QY 421 LGAPRYQHILGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHILGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480

QY 481 HYYEOTRGQSVSCPLPRGQARWOCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVSCPLPRGQARWOCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540

QY 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOQYFQSGSLGGQDLTMDGLV 600

DB 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOQYFQSGSLGGQDLTMDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRVKAINFENPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKAINFENPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLRREGIQSVVYTDALDSGRPHSRVFNENKSTRRTQTVGLTQTCTETKQLP 720
DB 661 STRDLRREGIQSVVYTDALDSGRPHSRVFNENKSTRRTQTVGLTQTCTETKQLP 720
QY 721 NCIEDPVPSPILVRLNFSLVGTPLSAFGLNRPVLAEDAQRLLFTALPFFKNCNDNICQDD 780
DB 721 NCIEDPVPSPILVRLNFSLVGTPLSAFGLNRPVLAEDAQRLLFTALPFFKNCNDNICQDD 780
QY 781 LSIITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTILQNORS 840
DB 781 LSIITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTILQNORS 840
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFQLELPVKYAYVMVTSKYSTKYNLTASNTSRVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTEFQLELPVKYAYVMVTSKYSTKYNLTASNTSRVMQHOYQVSN 960
QY 961 LGORSUPISLVFLVPVRLNQTWIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPV 1020
DB 961 LGORSUPISLVFLVPVRLNQTWIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPV 1020
QY 1021 VNCISIAVCQRIQCDIPFGIQEENATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVF 1080
DB 1021 VNCISIAVCQRIQCDIPFGIQEENATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVF 1080
QY 1081 TLLPGQGAFAVRSQTEKVPFEPVNPPLIVGSSVGLLLALLITAAALYKLGFFKQYKD 1140
DB 1081 TLLPGQGAFAVRSQTEKVPFEPVNPPLIVGSSVGLLLALLITAAALYKLGFFKQYKD 1140
QY 1141 MMSEGGPPGPAEPQ 1153
DB 1141 MMSEGGPPGPAEPQ 1153

RESULT 6
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX
AC AAR04136;
XX
XX 25-MAR-2003 (updated)
DT 07-SEP-1990 (first entry)
XX
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
XX synthetic.
XX
XX Location/Qualifiers
FH Key 86..88
FT Modified-site /*label= putative N-glycosylation site
FT Modified-site 240..242
FT /*label= putative N-glycosylation site
FT Modified-site 391..393
FT /*label= putative N-glycosylation site
FT Modified-site 469..471
FT /*label= putative N-glycosylation site
FT Modified-site 693..695
FT /*label= putative N-glycosylation site
FT Modified-site 697..699
FT /*label= putative N-glycosylation site

FT Modified-site 735..737 /*label= putative N-glycosylation site
 FT Modified-site 802..804 /*label= putative N-glycosylation site
 FT Modified-site 881..883 /*label= putative N-glycosylation site
 FT Modified-site 901..903 /*label= putative N-glycosylation site
 FT Modified-site 912..914 /*label= putative N-glycosylation site
 FT Modified-site 941..943 /*label= putative N-glycosylation site
 FT Modified-site 947..949 /*label= putative N-glycosylation site
 FT Modified-site 979..981 /*label= putative N-glycosylation site
 FT Modified-site 994..996 /*label= putative N-glycosylation site
 FT Modified-site 1022..1024 /*label= putative N-glycosylation site
 FT Modified-site 1045..1047 /*label= putative N-glycosylation site
 FT Modified-site 1051..1053 /*label= putative N-glycosylation site
 FT Modified-site 1076..1078 /*label= putative N-glycosylation site
 FT region 1..16 /*label= signal peptide
 FT region 1106..1134 /*label= putative transmembrane region

XX
 PN EP364690-A.

XX 25-APR-1990.

XX 17-AUG-1989; 89EP-0115159.

XX 23-AUG-1988; 88US-0235353;
 XX 09-MAR-1989; 89US-0321239

XX (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Springer TA, Corbi A;
 DR WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.

XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 XX
 PS Disclosure; Page ?; ?pp; English.

CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1153 AA;

Query Match 99.8%; Score 5946; DB 11; Length 1153;
 Best Local Similarity 99.9%;
 Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCGFLNLTENAMTFQENARGFQGVVQLQGSRVVVGAPQEIIVAANQR 60

DB 1 MALRVLLTALTLCGFLNLTENAMTFQENARGFQGVVQLQGSRVVVGAPQEIIVAANQR 60

QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVK 120

DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVK 120

QY 121 GLCFLFGSNLRQQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 DB 121 GLCFLFGSNLRQQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 QY 181 KSKTSLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
 DB 181 KSKTSLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSKSQEL 300
 DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSKSQEL 300
 QY 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
 DB 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
 QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTFRVDSMDNDAYLGYAAAIIILNRVQSLV 420
 DB 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTFRVDSMDNDAYLGYAAAIIILNRVQSLV 420
 QY 421 LGAPRYQHIGLVAMFRONTGHWESNANVKGTQIGAYFGASLCSDVDNSNGSTDVLVIGAP 480
 DB 421 LGAPRYQHIGLVAMFRONTGHWESNANVKGTQIGAYFGASLCSDVDNSNGSTDVLVIGAP 480
 QY 481 HYTEQTRGGQVSVCPFRGQARWQCDAVLYGEGQPMGRFGAALTVLGDVNGDKLTDVA 540
 DB 481 HYTEQTRGGQVSVCPFRGQARWQCDAVLYGEGQPMGRFGAALTVLGDVNGDKLTDVA 540
 QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYFGQSLSGGQDLTMDGLV 600
 DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYFGQSLSGGQDLTMDGLV 600
 QY 601 DLTVGAQGHVLLLRSPQVLRVKAIMEFNPREVAARVFNCECDQVVKGEAGVRVCLHVQK 660
 DB 601 DLTVGAQGHVLLLRSPQVLRVKAIMEFNPREVAARVFNCECDQVVKGEAGVRVCLHVQK 660
 QY 661 STRDLREGQIQSVVITYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP 720
 DB 661 STRDLREGQIQSVVITYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP 720
 QY 721 NCIEDPVPSPVLRNLFSLVGTPLSAFGNLRPVLAEADQRLFTALPPEKNGCNDNI CDD 780
 DB 721 NCIEDPVPSPVLRNLFSLVGTPLSAFGNLRPVLAEADQRLFTALPPEKNGCNDNI CDD 780
 QY 781 LSITFSEMSLDCLVVGPPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 DB 781 LSITFSEMSLDCLVVGPPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
 DB 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
 QY 901 NVTSENMPRTNKTFOLELPVKYAVVMVTSHGVSSTKYNLFTASNTSRVMQHOQVSN 960
 DB 901 NVTSENMPRTNKTFOLELPVKYAVVMVTSHGVSSTKYNLFTASNTSRVMQHOQVSN 960
 QY 961 LGQSLPLISLVFLVPRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
 DB 961 LGQSLPLISLVFLVPRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
 QY 1021 VNCISIAVCQRIQCDDIPFFGQIEEFNATLKGNLSDFWIKTSHNLLIIVSTAIEILFNDVSF 1080
 DB 1021 VNCISIAVCQRIQCDDIPFFGQIEEFNATLKGNLSDFWIKTSHNLLIIVSTAIEILFNDVSF 1080
 QY 1081 TLLPGQAFVRSQETETKVEPEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKRYQKD 1140
 DB 1081 TLLPGQAFVRSQETETKVEPEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKRYQKD 1140
 QY 1141 MMSEGGPPGAEPO 1153
 DB 1141 MMSEGGPPGAEPO 1153

Db 903 ENNTPRTSKTTFOLELPKYAVYTVVSSHEQFTKYNFSESEBESHVAMHYQVNNLQ 962
Qy 964 RSLPSLVLPVRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNC 1023
Db 963 RDLPSVINFVPELVNQBAMVDVEVSPQNPSCRSOKIAPASDPLAHIQNPVLD 1022
Qy 1024 SIACVORIQCDDIPFGIQEENATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLL 1083
Db 1023 SIAGCLRPRCDVPSPSQEELDFTLKGNSLFGWRQILQKKVSVVVAEITFDTSVYSQL 1082
Qy 1084 PGQAFVRSQETKVEPEVFNPLPLIVSSVGGLLILALITAAIYKLGFPKRYKQDMS 1143
Db 1083 PGOERFMAQTTTLEKYKVNPTPLIVGSSIGALLILALITAVLYKVGFPKRYKEMME 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 8
AAW65091
ID AAW65091 standard; Protein; 1163 AA.
XX AAW65091;
AC AAW65091;
XX 28-SEP-1998 (first entry)
DT Human Beta-integrin CD11c subunit protein.
DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
XX Homo sapiens.
OS US5728533-A.
XX 17-MAR-1998.
PD 07-JUN-1995; 95US-0485618.
XX 07-JUN-1995; 95US-0485618.
PR 23-DEC-1993; 93US-0172497.
PR 05-AUG-1994; 94US-0286889.
PR 21-DEC-1994; 94US-0362652.
XX (ICOS-) ICOS CORP.
PA Gallatin WM, Van DER VIEREN M;
PI WPI; 1998-206565/18.
DR Screening assay for modulators of integrin binding - using
XX immobilised or labelled alpha-d polypeptide, useful for, e.g.
PT treating type-1 diabetes
XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11c subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an
XX alpha-d binding partner, one of which is immobilised and the other of
XX which is labelled, in the presence of a test compound, and determining if
XX the compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
XX asthma, psoriasis, lung inflammation, acute respiratory distress
XX syndrome and rheumatoid arthritis.

XX SQ Sequence 1163 AA;
Query Match 58.4%; Score 3480; DB 19; Length 1163;
Best Local Similarity 61.1%; Pred. No. 7.5e-281;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
Qy 5 VLLLTALTLCCHGNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLY 64
Db 8 LLLFTALATSLGFLNLTDELTAFRVSAGFGSDVQVYANSVVVGAPQKIIAANQIGLY 67
Qy 65 QDYSTGSCBPIRLQVPPVEAVNMSLGLSLAATSPQLLACGPTVHOTCENTYVKGLCF 124
Db 68 QCGYSTGACBPIGLQVPPVEAVNMSLGLSLAATSPQLLACGPTVHOTCENTYVKGLCF 127
Qy 125 LFGSNLRQOPQKPEALRGCPQSDSDIAFLIDCGSII PHDFRPMKEFVSTMEQLKSK 184
Db 128 LLGPT--QLTQRLPVSQECPEQDQIVFLIDGSGSISSRNFMFMFVRAVISQFORPS 185
Qy 185 TLFSLMOYSEFRIHFTFEFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFTNTNG 244
Db 186 TQFSLMQFNSKFTQHTFEFRTSNPLSLASVHQLQFTYTATAIONVHRLFHASYG 245
Qy 245 ARKNAFKILVITDGEKFGDPLGYDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIA 304
Db 246 ARDAIKILIVITDGKKEGSDLDYKDVIPWADAAGIIRYAIQVGLAFQNRNSMKELNDIA 305
Qy 305 SKPPRDHVQVNNFEALKTIONQIREKI PAIEGTOTGSSSSSPEHMSQEGFSAITNGP 364
Db 306 SKPSQEHIFKVEDFDALQIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTDPG 365
Qy 365 LLGTGSGYDWAGVFLYTSKSTFNTMTRVDSMDNDAYLGAAAIILNRVQSLVIGAP 424
Db 366 VLGAUGSFTWSGAFLYPPNMSPTFINMSQENMDRSDYLGSTELAKWGVQSLVIGAP 425
Qy 425 RYQHIGLVAMFRQNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDLVLGAPHYE 484
Db 426 RYQHIGKAVIFIQVSRQWRMAEVIQTQIGSYFGASLCSVDVDTGSDTLVLGAPHYE 485
Qy 485 QTRGGQVSVCLPRGQARQWQCDVLYGEGQGWGRFGAALTVLGDVNGDKLTDVAIGAP 544
Db 486 QTRGGQVSVCLPRGWR-RWVCDVLYGEGQGWGRFGAALTVLGDVNGDKLTDVIGAP 544
Qy 545 GEEDNRGAVLPHGTSGSISPSHSORIASKLSPLQYFGOSLSCGODLTGMLVDLTV 604
Db 545 GEBENRGAVLPHGVLGSPISPSHSORIASKLSPLQYFGOSLSCGODLTGMLVDLTV 604
Qy 605 GAQGHVLLRSQVLRVKAIMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVOKSTRD 664
Db 605 GARGQVLLLRTPVLVWGVSMQFIPAEIPRSAFECEQVSEQTLVQSNTCLYIDKRSKN 664
Qy 665 RLREGQISVTVYDLALDSCGRPHSRVFNETKNSRRTQTVLGLTQTCETLKLQLPNCIE 724
Db 665 LLGSRDLQSSVTLDLALAPGLSPRAIFQETKNSRVRVVLGLKACHENFNLLLPSCVE 724
Qy 725 DPVSPVILRNFLSVLGTPLSAFGLNRPVLAEDAQRLFTALPPEKNCNNDNICODDISIT 784
Db 725 DSVIPIILRNFTLVGKPLLAFLRNPLAALAAQRYFTASLPPEKNCNNDNICODDISIT 784
Qy 785 FSPMSLDCLVGGPREFNVTVTVRNDESDSYRTQVTFPPPLDLSYKSVSTLQVRSORSW 844
Db 785 FSPFGLKSLLVGNLNLNNAEVMWNDESDSYRTQVTFPPPLDLSYKSVSTLQVRSORSW 844
Qy 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTC--CSAPVSGQSTWSTSCRNLHIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
Qy 905 ENNMPTNKTETEFOLELPVKYAVYTVVSSHEQFTKYNFSESEBESHVAMHYQVNNLQ 963
Db 903 ENNIPRTSKTTFOLELPKYAVYTVVSSHEQFTKYNFSESEBESHVAMHYQVNNLQ 962
Qy 964 RSLPSLVLPVRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNC 1023

Db 963 RDLPSINFWVPVELNQEAVMDVEVSHFQNPDLRCSSEKIAPPASDFLAHIQKNPVLDC 1022

QY 1024 STAVCORIQCDIPFFGIQEFNATLKGNSLFDWYIKTSHNHLIYSTAEILFENDSVFTLL 1083

Db 1023 SIAGCLRFCDVPFSFVSELDFTLKGNSLFGWQILOKKVSVSVASIIIFDTYSIQL 1082

QY 1084 PGGAFAVRGOTETKVPFFVNPPLPIVIGSSVGGILLALLALITAAALYKLGFFKQYKDMMS 1143

Db 1083 PGGAFAVRAQTITVLEKYKHNPPIPIVIGSSIGGLLLALLITAVLYKVGFKQYKEMME 1142

QY 1144 E 1144

Db 1143 E 1143

RESULT 9

AA007361

ID AAB07361 standard; Protein; 1163 AA.

XX

AC AAB07361;

XX

DT 17-JAN-2001 (first entry)

DE Human CD11c protein sequence.

XX

KW Human; macrophage infiltration inhibition; alpha d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11c.

XX

OS Homo sapiens.

XX

PN W0200029446-A1.

XX

PD 25-MAY-2000.

XX

PF 16-NOV-1999; 99WO-US27139.

XX

PR 16-NOV-1998; 98US-0193043.

PR 08-JUL-1999; 99US-0350259.

XX

PA (ICOS-) ICOS CORP.

XX

PI Gallatin MW, Van Der Vieren M;

XX

DR WPI; 2000-387751/33.

XX

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites -

XX

PS Example 5; Fig 1; 270pp; English.

XX

CC Integrins are a class of membrane-associated molecules that participate

CC in cellular adhesion. Integrins are made up of an alpha subunit and a

CC beta subunit. One class of human integrins are restricted to expression

CC in white blood cells and have a common beta2 subunit: the leukocyte

CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins

CC have an important role in immune and inflammatory responses. The present

CC protein sequence is the human integrin alpha subunit CD11c. This

CC sequence was used in an alignment to identify a novel beta2 integrin

CC alpha subunit: alpha d (AA060014 and AA07359). The present sequence has

CC approximately 66% identity to the protein sequence of alpha d. The

CC Alpha d gene and protein may be useful in therapy for diseases linked

CC to aberrant alpha d function e.g. Type I diabetes, atherosclerosis,

CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute

CC respiratory distress syndrome, rheumatoid arthritis and leukocyte

CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal

CC antibodies may be used in the inhibition of macrophage infiltration at

CC the site of a central nervous system injury. The monoclonal antibodies

CC can also be used to detect and diagnose Crohn's disease.

XX Sequence 1163 AA;

QY Query Match 58.4%; Score 3480; DB 21; Length 1163;

Best Local Similarity 61.1%; Pred. No. 7.5e-281;

Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLTALTCLCHGNLDTENAMTFQENARGFGQSVVQLOGSRVVGVGAPQEIIVAAHQSGLY 64

Db 8 LLLPTALATSLGFLNLTDELTAFRVDSAGFGDSVVQYANSVWVGAPQKIIAANQIGLY 67

QY 65 QCDYSTGCEPIRLQVPVEAVNMISGLSLAATTSPOLLACGPTVHQCSTNTYVKGICF 124

Db 68 QCGYSTGACEPIGLQVPPVEAVNMISGLSLASTTSPQLACGPTVHHCGRMYLTGICF 127

QY 125 LFGSNLRQOPKFPALRGCCPOEDSDIAFLIDGSGSIIPHDFRRMKFESTVMEOLKSK 184

Db 128 LIGPT--QLTQLPVSRQECPRQEQDIVFLIDGSGSISRNFATMNFVRAVISQFQPS 185

QY 185 TLFLSMQVSEEFRIHFTFKFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELNYNG 244

Db 186 TQFSLMQFSNKFQHTFTFEFRRTSNPLSLASVHQLQGFTYTATAIQNVHRLPHASYG 245

QY 245 ARKNAFKILVITDCEKFGDPLGYEDVIPADREGVIRVIGVGDAFSEKSRQELNTIA 304

Db 246 ARDAIKILIVITDKKGGDSLIDYKDVIPMAAAGIIRIAYGLVAFQNRNSWKELNDIA 305

QY 305 SKPPRDHVPQVNNFEALKTIONQLREKIPAIETGTTOTGSSSSFEHEMSQEGFAAITSNGP 364

Db 306 SKPSQEHIFKVEDFALKDIONQLKEKIPAIETGTTISSSSFELEMAQSGFAVTPDGP 365

QY 365 LLSTVGSVDWAGGVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAIIILNRVQSLVLAG 424

Db 366 VLGAVGSTWGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVSQSLVLAG 425

QY 425 RYQHIGLVAMPQNTGHWESNANVKGTOIGVFGASLCSVDVDSNGSTDLVLIGAPHYYE 484

Db 426 RYQHIGKAVIFTQVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDDTGDSTDLVLIGAPHYYE 485

QY 485 QTRGGQVSVCPPLRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAP 544

Db 486 QTRGGQVSVCPPLRGWR--RWCDVLYGEGHPWGRFGAALTIVLGVNGDKLTDVWIGAP 544

QY 545 GEDNRGAVYLFHGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTV 604

Db 545 GEBENRGAVYLFHGVLPSPISPSHSORIASGQSLSSRLQYFGQALSQGDITDGLVDLAV 604

QY 605 GAQGHVLLRSQPLRVKALMEFNPREVARNVECNQDVVKGKEAGEVRVCLHVOKSTR 664

Db 605 GARGQVLLKTRPVLWVGVSQMQFIPAEIPRSAFECEQVQVSEQTLVQSNICLYIDKRSKN 664

QY 665 RLREGQIQSVVYDYLALDSGRPHSRVAFNETKNSTRQVLTQVGLTQTCETLKLQLPNCIE 724

Db 665 LLGSRDLQSSVTLDLALAPGLSPRAIFQETKRSLSRVRLVGLKAHCENFLLPSCVE 724

QY 725 DPVSPVILRLNFSVLGTPLSAFQNLRPVLAEDAQRLFTALFPPEKNGNDNI CODLSIT 784

Db 725 DSVIPIILRLNFTLVGKPLLAFLRNLAPMLAALAAQRYFTASLPFEKNGADHICODNLGIS 784

QY 785 FSPMSLDCLVVGPRENFVTVTRNDEGDSYRQVTFPFFPLDLSYKRVSTLQNRORSW 844

Db 785 FSPFGLKSLVGNLELNAEVMYNDGDSYGTITITFHPAGUSYRVVAGSQKQGRSL 844

QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904

Db 845 HLITC--CSAPVSGQGTWSTSCRINHILFRGAQITFLATFDVSPKAVGLORULLIANVSS 902

QY 905 ENNMPRTNKTFFQLELPLVKYAVYVMTVSHGVSTKYLNFPTAS--ENTSRMHOHQVSNLQ 963

Db 903 ENNIPRTSKTIFQLELPLVKYAVYVMTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQ 962

QY 964 RSLPISLFLVPLVPLNQTIVDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVC 1023

Db	963	RDLPSVNFVWVPELNQEA	VMMDEVSH	PQNP	SRCS	SEKI	APPAS	FLAH	IKQNP	VLD	1082
Qy	1024	SIACVQRIQCDIPFGI	QEEFNATL	KGNSL	FDWIKT	SHNHL	LIVSTA	EILFN	DSVFTLL	1083	
Db	1023	SIACCLAFRCDDVP	SFSQBELD	FTLKG	NLSG	WRQILQ	KKVSV	VVAE	IFDTSV	SQL	1082
Qy	1084	PQGA	FVRSQ	TETK	VEP	EVNP	PLIV	GVSG	VGLLL	LALIT	AA
Db	1083	PGQEA	FVRSQ	TETK	VEP	EVNP	PLIV	GVSG	VGLLL	LALIT	AA
Qy	1144	E	1144								
Db	1143	E	1143								
RESULT	10										
ID	ABG61470										
ID	ABG61470	standard;	Protein;	1163	AA.						
AC	ABG61470;										
XX	XX										
DT	27-AUG-2002	(first entry)									
DE	Human Beta2 integrin alpha	CD11c	subunit.								
XX	Beta2 integrin;	alphaD	subunit;	CD11c	subunit;	CD11b	subunit;				
KW	LAD;	leukocyte	adhesion	deficiency;	inflammatory	response;	diabetes;				
KW	multiple	sclerosis;	arthritis;	graft	atherosclerosis;	neuroprotective;					
KW	inflammatory	bowel	disease;	Crohn's	disease;	ulcerative	colitis;				
KW	immune	complex	alveolitis;	leukaemia;	ICAM-1;	anti-inflammatory;					
KW	intracellular	cell	adhesion	molecule;	vascular	cell	adhesion	molecule;			
KW	locomotor	recovery;	locomotor	damage;	locomotor	impairment;					
XX	autonomic	dysfunction;	sensory	dysfunction;	spinal	cord	injury.				
OS	Homo sapiens.										
XX	WO200230980-A2.										
PN	18-APR-2002.										
PD	15-OCT-2001;	2001WO-US32059.									
PF	13-OCT-2000;	2000US-0688307.									
PR	(ICOS-)	ICOS CORP.									
PA	Gallatin	WM,	Van Der	Vieren	M;						
PI	WPI;	2002-463260/49.									
DR	Use of an anti-alpha-d	monoclonal	antibodies	for	promoting	locomotor					
PS	recovery, inhibiting	locomotor	damage,	limiting	locomotor	impairment,					
PS	or limiting	autonomic	and	sensory	dysfunction	following	spinal	cord			
XX	injury -										
XX	Example 5;	Page 194-198;	270pp;	English.							
XX	The invention	relates to	promoting	locomotor	recovery,	inhibiting					
CC	locomotor	damage, limiting	locomotor	impairment,	or limiting	autonomic					
CC	and sensory	dysfunction	following	spinal	cord	injury by	administering	an			
CC	anti-alpha-d	(Beta2	integrin	alpha2	subunit)	monoclonal	antibody	to	a		
CC	spinal	cord	injury	victim.	The	method	also	involves	the	use	of
CC	a ligand	selected	from	ICAM-R	or	VCAM-1	(intracellular	cell	adhesion		
CC	molecule,	vascular	cell	adhesion	molecule).	The	method	is	useful	for	
CC	promoting	locomotor	recovery,	inhibiting	locomotor	damage,	limiting				
CC	locomotor	impairment,	or	limiting	autonomic	and	sensory	dysfunction			
CC	following	spinal	cord	injury.	In	particular,	the	spinal	cord	injury	
CC	comprises	compression	of	the	spinal	cord.	The	antibodies	are	also	useful
CC	for	reducing	inflammation	at	the	site	of	a	central	nervous	system
CC	The	specification	also	details	the	identification	of	Beta2	integrin		
CC	alpha	CD11a	and	proteins,	for	use	in	raising	the	antibodies.	Beta2
CC	integrins	are	implicated	in	diseases	such	as	LAD	(leukocyte	adhesion	
CC	deficiency,	inflammatory	response,	diabetes,	multiple	sclerosis,					

Db 903 ENNIPRTSKTIFOLELKVYAVYVSSHEQFTKYNLFSESEKESHVAMHYQVNNLQ 962
Qy 964 RSLPISLVFLVPRLNQTIVDRPQVTFSENLSTCHTKERLPSHSDPLAELKAPVNC 1023
Db 963 RDLFVSINFVPELVNQEAVMDVEVSHQPNPSLRCSSEKIAPPASDFLAHQNPVLD 1022
Qy 1024 SIACVQRIQCDIPFGQEEFNATLKNLGFDMYIKTSHNHLIVSTAELFNDVFTLL 1083
Db 1023 SIACVQRIQCDIPFGQEEFNATLKNLGFDMYIKTSHNHLIVSTAELFNDVFTLL 1082
Qy 1084 PGQAFVRSQTEKVEPVPNPLIVGSGVGLLALLIATLALYKLGFFKQYKDMMS 1143
Db 1083 PGQAFVRSQTEKVEPVPNPLIVGSGVGLLALLIATLALYKLGFFKQYKDMMS 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 11

ABU07406
ID ABU07406 standard; Protein; 1163 AA.

XX AC ABU07406;

XX DT 28-JAN-2003 (first entry)

XX DE Protein differentially regulated in prostate cancer #9.

XX KW Prostate cancer; gene expression; differential regulation;
XX KW molecular marker; drug target; cancer detection; cancer diagnosis;
XX KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX OS Homo sapiens.

XX PN WO200281638-A2.

XX PD 17-OCT-2002.

XX PF 08-APR-2002; 2002MO-US10824.

XX PR 06-APR-2001; 2001US-281731P.

XX PR 06-APR-2001; 2001US-281732P.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PT Sun Z, Jay G;

XX PS WPI; 2003-058520/05.

XX PT Novel genes which are differentially regulated in prostate cancer,
XX PT useful for diagnosing prostate cancer in prostate tissue sample and
XX PT assessing therapeutic or preventive intervention in prostate cancer
XX PT patients -
XX PS Claim 1; Page 225-228; 416pp; English.

XX CC The invention describes genes (I) which are differentially regulated in
XX CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX CC sample comprising prostate tissue, which involves determining the number
XX CC of target genes which are differentially-regulated in the sample, where
XX CC the number is indicative of the probability that the sample comprises
XX CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX CC intervention in a subject having a prostate cancer, which involves
XX CC determining the expression levels in a sample comprising prostate tissue
XX CC of target genes which are differentially-regulated in prostate cancer.
XX CC Preferably, the expression levels of at least 10 genes are determined.
XX CC (I) is also useful for identifying agents that modulate a biological
XX CC activity of a polypeptide differentially-regulated in prostate cancer
XX CC cells, which involves contacting a polypeptide differentially-regulated
XX CC in prostate cancer cells with a test agent under conditions effective for
XX CC the test agent to modulate a biological activity of the polypeptide, and

CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.

XX Sequence 1163 AA;

Query Match 58.4%; Score 3478; DB 24; Length 1163;
Best Local Similarity 61.0%; Pred. No. 1.1e-280;
Matches 696; Conservative 136; Mismatches 303; Indels 6; Gaps 4;

Qy 5 VLLLTALTLCHEGFLNLTENAMTFQENARGQSGVQLOQGSRRVVGAPQETVAANQSGSLY 64
Db 8 LLLLTALTLCHEGFLNLTENAMTFQENARGQSGVQLOQGSRRVVGAPQETVAANQSGSLY 67
Qy 65 QCYSTGSCBPIRLQVPVEAVNMSLGLSLAATTPOLLACGPTVHOTCSENTYVKGLCF 124
Db 68 QCYSTGSCBPIRLQVPVEAVNMSLGLSLAATTPOLLACGPTVHOTCSENTYVKGLCF 127
Qy 125 LFGSNLRQOPKEPEALRGCPQEDSDIAFLIDCGSGTII PHDRMKFEFVSTVMEOLKSK 184
Db 128 LLGPT--QLTORLPVSRQCEPROEQDIVLIDGSGTSSSRNFATMNFVRAVISQFORPS 185
Qy 185 TLPSLMQYSEEFRIHFTFEFQNNPNRSLVKITQLLGRTHATATGIRKVVRELFINITNG 244
Db 186 TQSLMQFSNKFOHTLTFFEFRTSNPLSLASVHQLQGYTATATQNVVHRLPHASYG 245
Qy 245 ARKNAPKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIA 304
Db 246 ARDATKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIA 305
Qy 305 SKPRDRHVQNNFEALKTIONQLREKI FAIEGTOTGSSSSFEHMSQEGFSAITNSGP 364
Db 306 SKPSQEHIFKVEDFDALKDITQUREKIFPIEGTETSSSSFEHMSQEGFSAITNSGP 365
Qy 365 LLSTVGSYDWAQGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAP 424
Db 366 VLGNVGSFTWGGAFLYPPNMSPTFINMSQENVDMDRSLGYSTELALMKGVSQSLVGLAP 425
Qy 425 RYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGAGSLCSVDVDSNGSTDLVLGAPHYE 484
Db 426 RYQHTGKAVIFTQVSRQWRMKAEBVTGTIGSYFSPSLCSVDVDSNGSTDLVLGAPHYE 485
Qy 485 QTRGQVSVCPPLRGQBARWQCDALYGEQOPNGRFGAALTVLGDVNGDKLTDAVIGAP 544
Db 486 QTRGQVSVCPPLRGQBARWQCDALYGEQOPNGRFGAALTVLGDVNGDKLTDAVIGAP 544
Qy 545 GEENRGAVALFHCSTGSGISPSHSQRIAGSKLSPRQYFGQSLSGGQDLTMDGLVDLTV 604
Db 545 GEENRGAVALFHCSTGSGISPSHSQRIAGSKLSPRQYFGQSLSGGQDLTMDGLVDLTV 604
Qy 605 GAQGHVLLRSQPVLRVKAIMEFNPRVARNVFCNDQVVKGEAGEVRVCLHVQKSTRD 664
Db 605 GARGQVLLRTRPVLVWGVSMQFIPABIPRSAFECEQWSEQTLVQSNICLYIDKRSKN 664
Qy 665 RLREGQTSVVYDIALDSCPHSRVAFNETKNSRTROTQVLGLTQTCETLKLQLPNCIE 724
Db 665 LLGSRDLQSSVTLDALDPGRSLSPRATFQETKNRSLGRVRLGLKAHCENFNLLPSCVE 724

```
QY 725 DPVSPVLRNLSVLTPLSAGNLRPVLAEDAQRLLFTALFPFKNGNDNLCQDDLSIT 784
Db 725 DSVTPITRLNFTLVGKPLLAFLPNLRPLAADAQRYFTASLPFKNGGADHICQDNLGIS 784
QY 785 FSPMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPLDLSYRKVSTLQNSORSW 844
Db 785 FSPGGLKSLVGNLSLNAEVMVMNDGDSYGTITFSPAGLSYRYVAEGQKQQLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFFPENSEVTNITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTCDAPVG--SQGTWTSISCRINHILIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSS 902
QY 905 ENNMPTNKTBFQLELPVKYAVVMVTVSHGVSTKYLNFAS-ENTSRVMQHOYQVSNLQ 963
Db 903 ENNTPTSKTTFQLELPVKYAVVMVTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQ 962
QY 964 RSLPISLVLPVRLNOTWIDBPQVTFSENLSSTCHTKERLPSPHSDFLAELRKAPVNC 1023
Db 963 RDLPSVSNFVVELNQEAVMMDVEVSLPQNPSLRCSSEKIAGPASDFLAHTQKNPVLC 1022
QY 1024 STAVCORIOCDIPFFGIGQEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLL 1083
Db 1023 SIAGCLRFCDVPFSVQSELDFTLKNLSFGHVRQILQKKVSVSVAIITDTSVYSOL 1082
QY 1084 PGQAFVRSQETKVPFFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMS 1143
Db 1083 PGQAFMRAQTTVLEKYKHNPPLPLIVGSSIGLLLLALITAVLYKVGFFKRYKEMME 1142
QY 1144 E 1144
Db 1143 E 1143

RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX AC AAR78166;
XX DT 28-DEC-1995 (first entry)
XX DE Human beta-2 integrin alpha-d.
XX KW Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
XX KW inflammatory bowel disease; asthma.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 17..1108
XX FT /note= "extracellular domain"
XX FT Region 150..352
XX FT /note= "this region is homologous to the insert
XX FT common to Chila,b,c and may be a site for
XX FT interaction with ICAM family proteins"
XX FT Binding-site 465..474
XX FT /note= "putative cation binding site"
XX FT Binding-site 518..527
XX FT /note= "putative cation binding site"
XX FT Binding-site 592..600
XX FT /note= "putative cation binding site"
XX FT Region 1109..1128
XX FT /note= "transmembrane region"
XX FT Domain 1129..1161
XX FT /note= "cytoplasmic domain"
XX PN WO9517412-A1.
XX PD 29-JUN-1995.
XX XX
XX PF 21-DEC-1994; 94WO-US14832
XX XX
```

```
PR 05-AUG-1994; 94US-0286889.
PR 23-DEC-1993; 93US-0173497.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX DR WPI; 1995-240603/31.
XX DR N-PSDB; AAQ91712.
XX PT Alpha sub-unit polypeptide of human beta 2 integrin - used to
XX PT identify potential antiinflammatory agents, for the treatment of
XX PT graft arteriosclerosis, inflammatory bowel disease, asthma, etc.
XX PS Claim 7; Page 82-87; 172pp; English.
XX CC A probe based on a partial cDNA clone (given in AAQ91727) of canine
XX CC alpha-TM1 was used to screen a human spleen cDNA library to identify
XX CC clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.
XX CC The cDNA was manipulated to allow expression of recombinant alpha-d
XX CC subunit in COS and CHO cells.
XX SQ Sequence 1161 AA;
Query Match 58.0%; Score 3455; DB 16; Length 1161;
Best Local Similarity 59.9%; Pred. No. 9,1e-279;
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;
QY 5 VLLLTALTLCHGNLDTENAMTFQENARGFGQSVVQLOGSRVWVCAPOEIVAAORGSLY 64
Db 5 VLLLSVLASYGNLDFVEEPTIFQDAGGFGQSVVQFGRSLVVGAPLEVAANQTGRLY 64
QY 65 QCDYSGSCEPIRLQVPVEAVNMISLGLAATTSPPQLLACGPTVHOTCSENTYVKGLCF 124
Db 65 DCAATGNCQPIPLHIREAVNMISLGLTLAATNGSRLLACGPTLHRVCGENSYSKGSCL 124
QY 125 LFGSNLRQOPKFPALRGCPQEDSDIAFLIDGSGSIIPHDPRRKKEFVSTVMEQLKXK 184
Db 125 LLGSRW-EITQVDPATPECPHQEMDIVFLIDGSGSIDQDNFNQMGKFPVQAVMGQEGD 183
QY 185 TLFSLMYSSEPRIHFTFKEQNPNPRSLVKPITQLLGRHTATGIRKVVRELNITNG 244
Db 184 TLFALMQSNLLKIHFTFTQFTSPQSSLDVPIVQLKGLTFTATGILTVTVQLPHHKN 243
QY 245 ARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVGDAPRSEKSRQELNTIA 304
Db 244 ARKSAKKILVITDGQKYKDPLEYSDVIPAQEKAGIIRYVIGVGHAFQGTARQELNTIS 303
QY 305 SKPPRDHVPQVNNFEALKTIQNLREKIPAEIGTOTGSSSSSEHEMSQEGFSAAITNSGP 364
Db 304 SAPPQDHVPKVDNFAALGSIQKLOEKIYAVEGTOSRASSSFQHEMSQEGFSTALTM 363
QY 365 LLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLVGAAAAIILNRVQSLV 424
Db 364 FLGAVGSFWSGGLYPPNMSPTFINMSQENVMRDSYLGYSTELALWKGVQNLVLGAP 423
QY 425 RYQHIGLVAMPFRQNTGMESNANVAGTQIGAFYFASLCSVDVDSNGSTDLVLIGAPHY 484
Db 424 RYQHTGKAVIFTOVSRQWRKXAEVGTQIGSYFGLASLCSVDVDSGSTDLLIGAPH 483
QY 485 QTRGGQVSVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAP 544
Db 484 QTRGGQVSVCPPLPRGORVQWQCDVAVLRGEQHPWGRFGAALTIVLGDVNEKLDL 543
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOYFGQSLSGQDLTMDGLVDLTV 604
Db 544 GEQENRGAVYLFHGAESGISPSHSQRTASSQLSPRLOYFGQALSGQDLTQDGLMDLAV 603
QY 605 GAQGHVLLRSQPLVRKAIMENFREVARNVFECNDQVWKGKAGEVCLHVOKSTRD 664
Db 604 GARGQVLLRLSLPVLKVGAMRFPSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSS 663
QY 665 RLREGIQSVTVYDUALDSDGRPHSRVAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIE 724
```



```
Db 544 GEQNRGAVYLFHGAESGICSPSHSQRIASSQLSPRLQYFGQALSGGGDLTDGGLMDLAV 603
Qy 605 GAQGHVLLRSQVFLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRD 664
Db 604 GARGOVLLRLSLPVLKGVAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLD 663
Qy 665 RLREGOIQSVVTDIALDSGRPHSRVAFNETKNSRQTVGLGLTQTCETLKLQLPNCIE 724
Db 664 QL--GDIQSSVRFDLALDGRLTSAIFNETKNTPLTRKTLGLGIHCETLKLPLDCVE 721
Qy 725 DVPSPVLRLNLSVCTPLSAFNGNLRPVLAEADAQRLFTALRPFKNCGNDNICODLSIT 784
Db 722 DVSPFIHLNLSLVREPIPSQNLRPVLAVGSDQLFTASLDFEKNCGDGLCEGDLGV 781
Qy 785 F8FMSLDCLVWGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRW 844
Db 782 L8FSGLQTLTVGSSLELVITVMNAGEDSYGVTVSLYYPAGLSHRRVSGAQKQPHQSL 841
Qy 845 RIACESASTESVGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 842 RIACETV--PTEDEG--LRSSRCNVNHPHFHSGSNGTFIVTFDVSYKATLGDRLMLRASASS 899
Qy 905 ENNMPTNKTETOLELPVKYAVYVMVTVSHGVSTKYLN--TASENTSRVMQHOYQVSNLGO 963
Db 900 ENNKASSKATFQLELPVKYAVYTMISROESTKYFNFATSDEKXKXKAHRYRNNLSQ 959
Qy 964 RSLPISLVPLVRLNQTWIDRPQVTFSENLSSTCHTKERLPSHSDFLAEILRKAPVWNC 1023
Db 960 RLALISINFWPVLNGVAVWDMVMEAPSOQL--PCVSEKPPQSHDELTOISRSPMLDC 1017
Qy 1024 STAVCORIOCDIPFGIOBEFNATLKNLSFOWYIKTSHNHLLIVSTAEILPNDVSFTLL 1083
Db 1018 STADCLQFRCDVPFSFVQBELDFTLKNLSFGWVRETLLQKVLVSVVAEITPDTSVYSOL 1077
Qy 1084 PCOGAFVRSQETKVEPEVPNPLPLVIGSSVGGILLILALITAAALKYKLGFFKRYQKDWMS 1143
Db 1078 PCOEAFPMQAQMEWLEEDVYNAIPIIMGSSVGALELLALITATLYKLGFFKRYKEMLE 1137
Qy 1144 E 1144
Db 1138 D 1138

RESULT 14
AAW72825
ID AAW72825 standard; Protein; 1161 AA.
XX AC AAW72825;
XX DT 19-JAN-1999 (first entry)
XX DE Human alpha-d.
XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
KW immunoglobulin; fusion protein; binding molecule; antibody;
KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
KW acute respiratory distress syndrome.
XX OS Homo sapiens.
XX PN US5831029-A.
XX PD 03-NOV-1998.
XX PF 07-JUN-1995; 95US-0482293.
XX PR 07-JUN-1995; 95US-0482293.
XX PR 23-DEC-1993; 93US-0173497.
XX PR 05-AUG-1994; 94US-0286889.
XX PR 21-DEC-1994; 94US-0362652.
XX PA (ICOS-) ICOS CORP.

XX Gallatin WM, Van Der Vieren M;
XX WPI; 1998-609318/51.
XX N-PSDB; AAV67281.
XX Antibodies specific for beta-2 integrin alpha-subunit d - useful in
XX immunohistochemical analysis
XX Example 5; Column 61-66; 106pp; English.
XX The present sequence represents human alpha d. The present invention
XX describes: (1) an antibody that specifically binds alpha d, which
XX is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
XX antibody as in (1); (3) an anti-idiotypic antibody specific for the
XX monoclonal antibody of (2); (4) a hybridoma cell line producing the
XX monoclonal antibody of (2). Antibodies specific for alpha d can be
XX used in immunohistochemical analysis to localise alpha d to subcellular
XX compartments or individual cells within tissues. Substances that modulate
XX alpha d binding (which may include antibodies) can be used to treat
XX diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
XX inflammation, acute respiratory distress syndrome or arthritis.
XX Sequence 1161 AA;
XX Query Match 58.0%; Score 3455; DB 19; Length 1161;
XX Best Local Similarity 59.9%; Pred. No. 9.1e-279;
XX Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;
Qy 5 VLLLTALTLCGPNLDTENAMTFQENARCFQSVVQLQGSRRVVGAPQEIIVAAHQSGSLY 64
Db 5 VLLLSVLASYHGFNLDBEETIFQEDAGGFGQSVVQFQGSRLVVGAPLEVVAAHQSGRLY 64
Qy 65 QCDYSTGSCBPIRLQVPEAVNMSLGLSLAATSPQALLACGPTVHOCTSEYTVKGCFCF 124
Db 65 DCAATGMCQPIPHIRPEAVNMSLGLTAASTNGSLACGPTLHRVCGENSYSKSGSL 124
Qy 125 LFGNLRQOPKPFPEARLGCPEQSDIAFLIDGSGSIIPHPFRMKFVSTVMEQKKSK 184
Db 125 LLSRW--EIIQTVDPATPECPQEMDIVFLIDGSGSIDONDFNQMGFVQVAMQFEGTD 183
Qy 185 TLFSLMQYSEFRHFTFKFQNNPNRSLVKPTQLLGRTHRTATGIRKRVRELNIYNG 244
Db 184 TLFALMQYSNLLKHFTFTQRTSPSOQSLVDPIVLKGLTFTATGILTIVTQLPHKNG 243
Qy 245 ARKNAFKILVITDGEKFGDPLGVEDVITPEADREGVIRYVIGDGAFFSRKSEKRLNTIA 304
Db 244 ARKSAKILVITDQKYKOPLEYSDVIPQAEKAGITRYAIGVGHAFQGPATQRLNTIS 303
Qy 305 SKPPRDHVQVNFPEALKTIQNLREKIFAIEGTQTSSSSFEHEMSQEGFSAATISNGP 364
Db 304 SAPQDHFVKVDNFAALGSIQKQLEKIYAVEGTQSRASSFQHEMSQEGFSTALTMGL 363
Qy 365 LLSTVGSVDWAGGVPLVYTSKEKSTFINNTRVDSNDMDAYLGVAAILLRNRVQSLVLGAP 424
Db 364 FLGAVGSFWSGGAFLYPPNMSPTFINNSQSNVMDRDSYLGSTELALWKGVQNLVLGAP 423
Qy 425 RYOHIGLVAMFRONTGMHESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVE 484
Db 424 RYOHGTGKAVIFTQVSQRQWRKKAETVGTQIGSYFGASLCSVDVDSGDLTILIGAPHYVE 483
Qy 485 QTRGGQSVCPPLPRQARWOCDAVLVYGEQOPWGRFCAALTVLGDVNGDKLTDVAIGAP 544
Db 484 QTRGGQSVCPPLPRQARWOCDAVLVYGEQOPWGRFCAALTVLGDVNGDKLTDVAIGAP 543
Qy 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQALSGGGDLTDGGLMDLAV 604
Db 544 GEQNRGAVYLFHGAESGICSPSHSQRIASSQLSPRLQYFGQALSGGGDLTDGGLMDLAV 603
Qy 605 GAQGHVLLRSQVFLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRD 664
Db 604 GARGOVLLRLSLPVLKGVAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLD 663
```


Qy	545	GEENRGAVYLFHGTSGSISPSHSQRIAGSKLSPRLQYFCQSLSGGQDLTMDGLVDLTV	604
Db	544	GEENRGAVYLFHGTSGSISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAV	603
Qy	605	GAQGHVLLLRSQPLRVKAIWIEFNPREVARNVFCNDQVVKGEAGEVVRVCLHVQKSTRD	664
Db	604	GARGOVLLLRSLPVLKGVGVAMRFSEVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLD	663
Qy	665	RLREGQISVVTYDLALDSGRPHSRVFNETKNSRRQTVLGLTQTCETLKLQLPNCIE	724
Db	664	QL--GDIOSSVRFDLALDPGLTSRAIFNETKNPTLFRKTLGLGIHCETLKLPLDCVE	721
Qy	725	DPVSPVLRLNFSLVGTPLSAFNLRPVLAEDAQRFTALFPFKKNCNDNICODDLSIT	784
Db	722	DWSPILHLNFSLVREIPSPQNLRPVLAVGSQDLFTASLPFKKNCQDGLCEGLGVT	781
Qy	785	FSFMSLDCLVGGPREFNVTVRNDGEDSYRTQVTFPFLDLSYRKYSTLQNRQSRWS	844
Db	782	LSFSGLOTLTVGSSLELNVIVTMNAGEDSYGTVVSLYPAGLSHRRVSGAQKQPHOSAL	841
Qy	845	RLACESASSTEVSGALKSTSCSINHPIEPENSEVTFNITFDVDSKASLGNKLLKANVTS	904
Db	842	PLACETV-PTDEG-LRSRCSVNHPIPHESNGTFIVTFDVSYKATLGDRLMLRASASS	899
Qy	905	ENNMPTNKTFFQLELPVKYAVYVMVTSHGYSTKYNF-TASENSTRVMQHOYQVSNLQ	963
Db	900	ENNKASSKATFQLELPVKYAVYTMISQEESTKYFNFATSDEKQKKEAHRVYVNNLSQ	959
Qy	964	RLPISLVLPVRLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVC	1023
Db	960	RDLAISINFWPVLINGVAVMDVMEAPSQL--PCVSEKPPQHSDFLTQISRSPMLDC	1017
Qy	1024	SIACQRIQCDDIPFGIOEFENATLKGNLSDWIKTSHNHLIVSTAEILFNDSVFTLL	1083
Db	1018	SIADCLQRCQVPSFVSQBELDFTLKGNLSCFWVRETQKKVLVVSVAEITFDTSVYSQL	1077
Qy	1084	PGQAFVRSQTEKVEPFPVNPPLIYGSSVGGLLLALITAALYKLGFFKRYQKMMMS	1143
Db	1078	PGQAFMEAQMMEVLEDEVYNAIFIINGSSVGALLLALITATLYKLGFFKRYKEMLE	1137
Qy	1144	E	1144
Db	1138	D	1138

Search completed: November 25, 2003, 14:16:22
Job time : 40.8086 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:20:45 ; Search time 26.4941 Seconds
(without alignments)
8026.803 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLTLTALTLCHGFNLD.....FKRQKDMSEGPGCAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 18443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5956	100.0	1153	9	US-09-350-259-3
2	5956	100.0	1153	11	US-09-902-481A-1
3	5956	100.0	1153	11	US-09-891-943-3
4	5956	100.0	1153	15	US-10-144-259-30
5	5956	100.0	1153	15	US-10-207-655-176
6	5940.5	99.7	1152	10	US-09-945-265-4
7	5868	98.5	1137	11	US-09-902-481A-6
8	5862	98.4	1137	11	US-09-902-481A-5
9	5852	98.3	1137	11	US-09-902-481A-4
10	5839	98.0	1137	11	US-09-902-481A-3
11	3503	58.8	1163	12	US-10-116-275-204
12	3480	58.4	1163	9	US-09-350-259-4
13	3480	58.4	1163	11	US-09-891-943-4
14	3455	58.0	1161	9	US-09-350-259-2
15	3455	58.0	1161	11	US-09-891-943-2

16	3439.5	57.7	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3439.5	57.7	1161	11	US-09-891-943-99	Sequence 99, Appl
18	3270.5	54.9	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3270.5	54.9	1161	11	US-09-891-943-55	Sequence 55, Appl
20	3264	54.8	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3264	54.8	1161	11	US-09-891-943-53	Sequence 53, Appl
22	3251.5	54.6	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3251.5	54.6	1151	11	US-09-891-943-37	Sequence 37, Appl
24	3243	54.4	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3243	54.4	1155	11	US-09-891-943-46	Sequence 46, Appl
26	1563	26.2	1170	10	US-09-945-265-2	Sequence 2, Appl
27	1229.5	20.6	494	9	US-09-350-259-103	Sequence 103, App
28	1229.5	20.6	494	11	US-09-891-943-103	Sequence 103, App
29	1189.5	20.0	413	9	US-09-350-259-101	Sequence 101, App
30	1189.5	20.0	413	11	US-09-891-943-101	Sequence 101, App
31	1163	19.5	1179	12	US-10-177-550-2	Sequence 2, Appl
32	1108.5	18.6	1151	11	US-09-984-130-103	Sequence 103, App
33	1108.5	18.6	1151	12	US-09-836-353A-103	Sequence 103, App
34	1102.5	18.5	1189	11	US-09-984-130-35	Sequence 35, Appl
35	1102.5	18.5	1189	12	US-09-836-353A-35	Sequence 35, Appl
36	1087.5	18.3	1181	15	US-10-160-354-2	Sequence 2, Appl
37	1042	17.5	216	12	US-09-795-872-5	Sequence 5, Appl
38	1036	17.4	240	15	US-10-102-806-453	Sequence 453, App
39	987	16.6	1034	11	US-09-984-130-43	Sequence 43, Appl
40	987	16.6	1034	12	US-09-836-353A-43	Sequence 43, Appl
41	979	16.4	191	11	US-09-805-354-1	Sequence 1, Appl
42	979	16.4	191	15	US-10-144-259-1	Sequence 1, Appl
43	875	14.7	707	9	US-09-764-870-313	Sequence 313, App
44	875	14.7	707	15	US-10-125-540-313	Sequence 313, App
45	850	14.3	223	12	US-10-002-631C-197	Sequence 197, App

ALIGNMENTS

RESULT 1

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 100.0%; Score 5956; DB 9; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLTLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVANOR 60
DB 1 MALRVLTLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVANOR 60
QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMISLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120

Db 61 GSYLQCDYSGCEPIRLQVPVEAVNMSGLSLAATTSPOLLACGPTVHTCSENTYVK 120
QY 121 GLCLFSGNLQQPOKFPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Db 121 GLCLFSGNLQQPOKFPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
QY 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Db 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
QY 481 HYYEOTRGQGVSVCPPLRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQGVSVCPPLRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
QY 601 DLTUGAQGHVLLRSQVLRVKAIMEFNPREVARNVFECDQVVKGKAGEVRVCLHVQK 660
Db 601 DLTUGAQGHVLLRSQVLRVKAIMEFNPREVARNVFECDQVVKGKAGEVRVCLHVQK 660
QY 661 STRDLRREGIOISVVTYDLALDSGRPHSRAVENETKNSRRQTQVLGTOTCETLKLQLP 720
Db 661 STRDLRREGIOISVVTYDLALDSGRPHSRAVENETKNSRRQTQVLGTOTCETLKLQLP 720
QY 721 NCIEDPVSPILVRLNFSLVGTPLSAFNGNRPVLAEDAORLFTALPFPFKNCNDNICQDD 780
Db 721 NCIEDPVSPILVRLNFSLVGTPLSAFNGNRPVLAEDAORLFTALPFPFKNCNDNICQDD 780
QY 781 LSITFSFMSLCLVVGGRPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRS 840
Db 781 LSITFSFMSLCLVVGGRPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRS 840
QY 841 QRSWRLACESASTEVSGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACESASTEVSGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNKTBFQELPVLKYAVYVMTVSHGVSTKYLNFTASENTSRVNHQYQVSN 960
Db 901 NVTSENNMPRTNKTBFQELPVLKYAVYVMTVSHGVSTKYLNFTASENTSRVNHQYQVSN 960
QY 961 LGQRLSPLSLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LGQRLSPLSLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSF 1080
Db 1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSF 1080
QY 1081 TLLPGQAFVRSQETETKVEPFVPPNPLPLIVGSSVGGLLLLALITAAALYKLGFFFRQYKD 1140
Db 1081 TLLPGQAFVRSQETETKVEPFVPPNPLPLIVGSSVGGLLLLALITAAALYKLGFFFRQYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
|||||

Db 1141 MMSEGGPPGAEPQ 1153
RESULT 2
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/REF/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 100.0%; Score 5956; DB 11; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLTENAMTFQENARGQSVQLOQSRVVVVGAPQEIIVANQR 60
Db 1 MALRVLLLTALTLCGHNLTENAMTFQENARGQSVQLOQSRVVVVGAPQEIIVANQR 60
QY 61 GSYLQCDYSGCEPIRLQVPVEAVNMSGLSLAATTSPOLLACGPTVHTCSENTYVK 120
Db 61 GSYLQCDYSGCEPIRLQVPVEAVNMSGLSLAATTSPOLLACGPTVHTCSENTYVK 120
QY 121 GLCLFSGNLQQPOKFPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Db 121 GLCLFSGNLQQPOKFPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
QY 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Db 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
QY 481 HYYEOTRGQGVSVCPPLRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQGVSVCPPLRGQARWQCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600

```
QY 601 DLTGAGQHVLLRSQPLVRKAIEMFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTGAGQHVLLRSQPLVRKAIEMFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVNETKNSRRQTVGLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVNETKNSRRQTVGLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSARGNLRPVLAEDAQRLLFTALPFEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSARGNLRPVLAEDAQRLLFTALPFEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQSN 960
DB 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQSN 960
QY 961 LQORSPLISLVFLVPLRNLQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
DB 961 LQORSPLISLVFLVPLRNLQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
QY 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAILFNDVSF 1080
DB 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAILFNDVSF 1080
QY 1081 TLLPGQGA FVRSQTEKVPPEFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKD 1140
DB 1081 TLLPGQGA FVRSQTEKVPPEFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153
```

RESULT 3

```
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3
```

```
Query Match 100.0%; Score 5956; DB 11; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAHQ 60
DB 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAHQ 60
QY 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTCSENTYVK 120
DB 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTCSENTYVK 120
QY 121 GLCFLGSLNLRQOQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRBMKFSVSTMESOL 180
DB 121 GLCFLGSLNLRQOQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRBMKFSVSTMESOL 180
QY 181 KKSKTFLSLMOYSEEFRIHFTFKFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSKTFLSLMOYSEEFRIHFTFKFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300
QY 301 NTIASKPRDRHVQVNNFEALKTIQNLREKI FAEGTQTGSSSSFEHMSOEGFSAIT 360
DB 301 NTIASKPRDRHVQVNNFEALKTIQNLREKI FAEGTQTGSSSSFEHMSOEGFSAIT 360
QY 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
DB 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASICSDVDNSGSDTLVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASICSDVDNSGSDTLVLIGAP 480
QY 481 HYYETRQGOVSVCPLPRGQARWQCDAVLYGEOQPGWRFGAALTVLGDVNGDKLTQVA 540
DB 481 HYYETRQGOVSVCPLPRGQARWQCDAVLYGEOQPGWRFGAALTVLGDVNGDKLTQVA 540
QY 541 IGAPCEENRGAVALYFHGTSGSIGSPSHSQRISAGSLSPRLQYFCQSLSGGQDLTMDGLV 600
DB 541 IGAPCEENRGAVALYFHGTSGSIGSPSHSQRISAGSLSPRLQYFCQSLSGGQDLTMDGLV 600
QY 601 DLTGAGQHVLLRSQPLVRKAIEMFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTGAGQHVLLRSQPLVRKAIEMFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVNETKNSRRQTVGLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVNETKNSRRQTVGLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSARGNLRPVLAEDAQRLLFTALPFEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSARGNLRPVLAEDAQRLLFTALPFEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQSN 960
DB 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQSN 960
QY 961 LQORSPLISLVFLVPLRNLQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
DB 961 LQORSPLISLVFLVPLRNLQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
QY 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAILFNDVSF 1080
DB 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSHNHLIIVSTAILFNDVSF 1080
QY 1081 TLLPGQGA FVRSQTEKVPPEFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKD 1140
```

```
Db 1081 TLLPGQAFVRSQTEKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFPKQYKD 1140
Qy 1141 MMSEGGPPGAPQ 1153
Db 1141 MMSEGGPPGAPQ 1153

RESULT 4
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match 100.0%; Score 5956; DB 15; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTALCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIWAANQR 60
Db 1 MALRVLLLTALTALCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIWAANQR 60

Qy 61 GSLYQCDYSTGSCPEIRLQVPEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120
Db 61 GSLYQCDYSTGSCPEIRLQVPEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120

Qy 121 GLCFPLGSLNLRQOPKFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQL 180
Db 121 GLCFPLGSLNLRQOPKFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQL 180

Qy 181 KKSFTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSFTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240

Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300

Qy 301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGESAIT 360
Db 301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGESAIT 360

Qy 361 SNGPLLSLVGSDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVQSLV 420
Db 361 SNGPLLSLVGSDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVQSLV 420

Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLIGAP 480

Qy 481 HYYETRGQSVSCPLPRGQRARWQCDVLYGEOQOPWGRGAALTVILGDVNGDKLTDVA 540
Db 481 HYYETRGQSVSCPLPRGQRARWQCDVLYGEOQOPWGRGAALTVILGDVNGDKLTDVA 540

Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
```

```
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAGQHVLRLRSQPLRVKAIMFENPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
Db 601 DLTVGAGQHVLRLRSQPLRVKAIMFENPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
Qy 661 STRDLREGQIQSWVTYDIALDSGRPHSRVAFNETKNSTRROTQVGLGLTQTCETLKLQLP 720
Db 661 STRDLREGQIQSWVTYDIALDSGRPHSRVAFNETKNSTRROTQVGLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPIVLRNLSVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDD 780
Db 721 NCIEDPVSPIVLRNLSVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDD 780
Qy 781 LSITFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLVSRKVKSTLQNR 840
Db 781 LSITFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLVSRKVKSTLQNR 840
Qy 841 QRSWRLACESASSTEVSAGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACESASSTEVSAGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNMPTNKTEFOLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHQVSN 960
Db 901 NVTSENNMPTNKTEFOLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHQVSN 960
Qy 961 LGORSUPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LGORSUPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVF 1080
Qy 1081 TLLPGQAFVRSQTEKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFPKQYKD 1140
Db 1081 TLLPGQAFVRSQTEKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFPKQYKD 1140
Qy 1141 MMSEGGPPGAPQ 1153
Db 1141 MMSEGGPPGAPQ 1153

RESULT 5
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 100.0%; Score 5956; DB 15; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTALCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIWAANQR 60
Db 1 MALRVLLLTALTALCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIWAANQR 60

Qy 61 GSLYQCDYSTGSCPEIRLQVPEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120
```

```
Db 61 GSYLQCDYSTGSCPTIRLQVPVBAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFLFGSNLRQOPQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
Qy 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Qy 301 NTIASKPPRDHVQVNNFALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAIT 360
Db 301 NTIASKPPRDHVQVNNFALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAIT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILRNRVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILRNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYVEOTRGGQSVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYVEOTRGGQSVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSRLOVFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSRLOVFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
Qy 661 STRDLREGOIQSVVTVYDALDSGRPHSRAVFNKTNSTRTOVNLGLTQTCETLKLQIP 720
Db 661 STRDLREGOIQSVVTVYDALDSGRPHSRAVFNKTNSTRTOVNLGLTQTCETLKLQIP 720
Qy 721 NCIEDPVSPVLRLNFSVGTPLSAFGNLRPVLAEADAQRLFTALFPFEKNCNDNICQDD 780
Db 721 NCIEDPVSPVLRLNFSVGTPLSAFGNLRPVLAEADAQRLFTALFPFEKNCNDNICQDD 780
Qy 781 LSITFSFMSLDCLVVGGRPEFNTVTVRNDGEDSYRTQVTFPPPLDLRYKRVSTLQNRS 840
Db 781 LSITFSFMSLDCLVVGGRPEFNTVTVRNDGEDSYRTQVTFPPPLDLRYKRVSTLQNRS 840
Qy 841 ORSMWLACASASTEVSGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKA 900
Db 841 ORSMWLACASASTEVSGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRTNKTETQELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
Db 901 NVTSENMPRTNKTETQELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
Qy 961 LQORSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020
Db 961 LQORSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020
Qy 1021 VNCIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVF 1080
Db 1021 VNCIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVF 1080
Qy 1081 TLLPGOGAFVRSOTETKVEPFPNPLPLTVGSSVGGLLLLALITAAALYKLGFFKKQYKD 1140
Db 1081 TLLPGOGAFVRSOTETKVEPFPNPLPLTVGSSVGGLLLLALITAAALYKLGFFKKQYKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153
```

RESULT 6

```
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

```
Query Match 99.7%; Score 5940.5; DB 10; Length 1152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEI1VAANOR 60
Db 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEI1VAANOR 60
Qy 61 GSYLQCDYSTGSCPTIRLQVPVBAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
Db 61 GSYLQCDYSTGSCPTIRLQVPVBAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFLFGSNLRQOPQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
Qy 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Qy 301 NTIASKPPRDHVQVNNFALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAIT 360
Db 301 NTIASKPPRDHVQVNNFALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAIT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILRNRVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILRNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYVEOTRGGQSVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYVEOTRGGQSVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSRLOVFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSRLOVFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
```


Qy	661	STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLP	720
Db	660	STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLP	719
Qy	721	NCIEDPVSPIVLRNLSLVGTPLSAFGNLRPVLAEADQRLFTALPPFEKNCNDNI	780
Db	720	NCIEDPVSPIVLRNLSLVGTPLSAFGNLRPVLAEADQRLFTALPPFEKNCNDNI	779
Qy	781	LSITFSFMSLCLVVGGRPRENVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR	840
Db	780	LSITFSFMSLCLVVGGRPRENVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR	839
Qy	841	QSRWLACASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA	900
Db	840	QSRWLACASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA	899
Qy	901	NVTSENNMPTNKTEFQLELPLVKYAVVMVTSHGVSSTKYNFTASENTRVMQHOYVSN	960
Db	900	NVTSENNMPTNKTEFQLELPLVKYAVVMVTSHGVSSTKYNFTASENTRVMQHOYVSN	959
Qy	961	LGORSPLSLVPLVRLNQTIVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV	1020
Db	960	LGORSPLSLVPLVRLNQTIVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV	1019
Qy	1021	VNCSTAVCORIQCDIPFFGIGOEFEFNATLKGNSLFDWYIKTSHNLLIIVSTAEIIFNDSVF	1080
Db	1020	VNCSTAVCORIQCDIPFFGIGOEFEFNATLKGNSLFDWYIKTSHNLLIIVSTAEIIFNDSVF	1079
Qy	1081	TLLPQOGAFVRSQTKVPEVNPPLPLIVGSSVGGLLLALITAAALYKLGPFKQYKD	1140
Db	1080	TLLPQOGAFVRSQTKVPEVNPPLPLIVGSSVGGLLLALITAAALYKLGPFKQYKD	1139
Qy	1141	MMSEGGPPGAEPQ 1153	
Db	1140	MMSEGGPPGAEPQ 1152	
RESULT 7			
US-09-902-481A-6			
; Sequence 6, Application US/09902481A			
; Publication No. US2003005440A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Shifman, Julia			
; APPLICANT: Mayo, Stephen			
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY			
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK			
; CURRENT APPLICATION NUMBER: US/09/902,481A			
; PRIOR FILING DATE: 2001-07-09			
; PRIOR FILING DATE: 2000-07-09			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; LENGTH: 1137			
; TYPE: PRT			
; ORGANISM: Artificial sequence			
; FEATURE:			
; OTHER INFORMATION: synthetic			
US-09-902-481A-6			
Query Match 98.5%; Score 5868; DB 11; Length 1137;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;			
Qy	17	FNLDTENAMTFENARGFQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCBPI	76
Db	1	FNLDTENAMTFENARGFQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCBPI	60
Qy	77	RLQVPVEAVNMSLGLSLAATTGPPQLLACGPTVHQTCSNTYVVGKLCFLFGSNLRQPOK	136
Db	61	RLQVPVEAVNMSLGLSLAATTGPPQLLACGPTVHQTCSNTYVVGKLCFLFGSNLRQPOK	120

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A

181 RIHFTKFEQNNPNRSLIKPIITQLLGRTHATGLRKVRVRELNTNGARKNAFKILFLL 240
QY 257 TDEKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTTASPPRDHVFQVN 316
Db 241 TDGKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
QY 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 376
Db 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 360
QY 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 436
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420
QY 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCP 496
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCP 480
QY 497 PRGORARWQCDAVLYGSGQSPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYL 556
Db 481 PRGORARWQCDAVLYGSGQSPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYL 540
QY 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 616
Db 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 600
QY 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Db 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 677 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNF 736
Db 661 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNF 720
QY 737 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLV 796
Db 721 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLV 780
QY 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRORSWRLACESASSTEV 856
Db 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRORSWRLACESASSTEV 840
QY 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
Db 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
QY 917 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMQHYQVSNLQORSLSPLSLVFLPV 976
Db 901 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMQHYQVSNLQORSLSPLSLVFLPV 960
QY 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
Db 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
QY 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAGAFVRSQ 1096
Db 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAGAFVRSQ 1080
QY 1097 KVEPEVPNPPLIIVGSSVGGLLLLALITAALYKLGFPRQYKQMMSEGGPPGAEPP 1153
Db 1081 KVEPEVPNPPLIIVGSSVGGLLLLALITAALYKLGFPRQYKQMMSEGGPPGAEPP 1137

RESULT 10

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimadzu, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

FILE REFERENCE: A-70586-1/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 1137
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 98.0%; Score 5839; DB 11; Length 1137;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVAAANQSGLYOCYSTGCEPI 76
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVAAANQSGLYOCYSTGCEPI 60
QY 77 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFGLSGNLROPOK 136
Db 61 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFGLSGNLROPOK 120
QY 137 FPBALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQLKSKTFLSLMQYSEEF 196
Db 121 FPBALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
QY 197 RIHFTKFEQNNPNRSLIKPIITQLLGRTHATGLRKVRVRELNTNGARKNAFKILFLL 256
Db 181 RIHFTKFEQNNPNRSLIKPIITQLLGRTHATGLRKVRVRELNTNGARKNAFKILFLL 240
QY 257 TDEKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTTASPPRDHVFQVN 316
Db 241 TDEKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTTASPPRDHVFQVN 300
QY 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 376
Db 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 360
QY 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 436
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420
QY 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCP 496
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCP 480
QY 497 PRGORARWQCDAVLYGSGQSPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYL 556
Db 481 PRGORARWQCDAVLYGSGQSPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYL 540
QY 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 616
Db 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 600
QY 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Db 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 677 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNF 736
Db 661 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNF 720
QY 737 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLV 796
Db 721 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSPMSLDCLV 780
QY 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRORSWRLACESASSTEV 856


```

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQSGSLY 64
DB 8 LLLFTALATSLGFLNLTDELTAFRVDSAGFGDSVVQYANSVVVGAPQKIITAAQIGGLY 67

QY 65 QCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACEPIGLQVPPEAVNMSLGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCF 127

QY 125 LFGSNLRQOPKQFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSK 184
DB 128 LLGPT--QLTQRLPVSROCEPRQEQDIVFLIDGSGSISSRNPFATMNFVRAVISQFORPS 185

QY 185 TLFSLMQVSEFRTHFTKFEQNNPNPRSLVKPITOLLGRTHATGIRKVVRELFNTNG 244
DB 186 TQFSLMQPSNKFQTHFTFEFRRTSNPLSLASVHQLQGFYTTATATQNVVHRLFHASYG 245

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 785 FSFPGLKSLVGNLNLNAEVMVMNDGEDSYGTTITFSHPAGLSYRYVABGQKQGLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSVTNIITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTC--CSAPVGSQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
QY 905 ENNMPTNKTEFQLELPVKYAVYVMVTSYHGVSTKYLNFTAS--ENTSRVMOHQYOVSNLQ 963
DB 903 ENNIPRTSKTIFQLELPVKYAVYVSSHEQTKYLNFSSEESKESHVAMHRYOVNVLQ 962
QY 964 RSLPISLVFLVPLVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNC 1023
DB 963 RDLFVSINFVWPVELNQEAVMVMDEVSHQPNPQLSRCSSEKIAIPASDFLAHIQKNPVLDC 1022
QY 1024 SIACVQRIQCDIPFPGIQQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLL 1083
DB 1023 SIACCLAFRCDFVSFSQVEELDTLKNLSFGVVRQILQKKVSVSVAEIIIDTYSVSQL 1082
QY 1084 PGQAFVRSQETKVEPEFVNPPLPLIVGSSVGLLLALITAAALYKGLFFKQYKQDMMS 1143
DB 1083 PGQAFVRAQTITVLEKYKVHNPILVIGSSIGGLLLALITAVLYKVGFVKQYKEMME 1142
QY 1144 E 1144
DB 1143 E 1143

RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943
; Publication No. US2003007728A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.4%; Score 3480; DB 11; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQSGSLY 64
DB 8 LLLFTALATSLGFLNLTDELTAFRVDSAGFGDSVVQYANSVVVGAPQKIITAAQIGGLY 67

QY 65 QCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACEPIGLQVPPEAVNMSLGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCF 127

QY 125 LFGSNLRQOPKQFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSK 184
DB 128 LLGPT--QLTQRLPVSROCEPRQEQDIVFLIDGSGSISSRNPFATMNFVRAVISQFORPS 185

QY 185 TLFSLMQVSEFRTHFTKFEQNNPNPRSLVKPITOLLGRTHATGIRKVVRELFNTNG 244
DB 186 TQFSLMQPSNKFQTHFTFEFRRTSNPLSLASVHQLQGFYTTATATQNVVHRLFHASYG 245
```

[illegible]

RESULT 14

RESULT 14
US-09-350-259-2

US-09-350-259-2 : Sequence 2: Application IIS/09350259

sequence 2, APPLICATION US
Patent No. US20020062008A1

; PATENT NO. US2002006
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.

APPLICANT: Van der Vieren, Monica

Qy 725 DPVSPVLRNLSVGLVPLSAFNGNLRPVLAEDAQRFTALFPFPEKNCNDNI CODDLSIT 784
 Db 722 DWSPPIILHLNLSLVREPISQNLRLPVLA VGSQDLFTASLPFPEKNCQDGLCEGLGVT 781
 Qy 785 FSNLSCLVVGVPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRORSW 844
 Db 782 LSFSGLOTLTVGSSLELNVIITVMNAGEDSVTVVSLYPAGLSHRRVSGNQKOPHQSAL 841
 Qy 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
 Db 842 RLACETV-PTDEG-LRSSRCSVNHPIFHEGSGNGTFIVTDFVSYKATLGDRLMRASASS 899
 Qy 905 ENNMPRTNKTFFOLELPVKYAVYVMVTSYHGVSTKYLPF-TASENTRVMQHOYQVSNLQ 963
 Db 900 ENNKASSSKATFOLELPVKYAVYTMISROEESTKYFNATSDKKMKEAEHRYVNNLSQ 959
 Qy 964 RSLPISLVLPVRLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVC 1023
 Db 960 RDLAISINFWPVLNGVAVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDC 1017
 Qy 1024 STIAVCORIQCDDIPFGIQQEENATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLL 1083
 Db 1018 SIADCLQPRCDVPSPSQEELDFTLKGNSLPGWVRETLQKKVLVVSAEITFTDTSVSQL 1077
 Qy 1084 PQGAFVRSOTETKVEPEVNPPLIVGSSVGLLALLALITAAIYKLGFFKROYKDMMS 1143
 Db 1078 PQEAFMRAQMEMVLEEDVYNAIPIIMGSSVGALLLALLITATLYKLGFFKRYKEMLE 1137
 Qy 1144 E 1144
 Db 1138 D 1138

RESULT 15

US-09-891-943-2
 ; Sequence 2, Application US/09891943
 ; Publication No. US2003007728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.
 ; APPLICANT: Van der Vieren, Monica
 ; TITLE OF INVENTION: No. US2003007728A1el Human 2
 ; FILE REFERENCE: 27866/35004
 ; CURRENT APPLICATION NUMBER: US/09/891,943
 ; CURRENT FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 09/193,043
 ; PRIOR FILING DATE: 1998-11-16
 ; PRIOR APPLICATION NUMBER: 08/286,889
 ; PRIOR FILING DATE: 1994-08-05
 ; PRIOR APPLICATION NUMBER: 08/362,652
 ; PRIOR FILING DATE: 1994-12-21
 ; PRIOR APPLICATION NUMBER: 08/943,363
 ; PRIOR FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1161
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-891-943-2

Query Match 58.0%; Score 3455; DB 11; Length 1161;
 Best Local Similarity 59.9%; Pred. No. 2.9e-308;
 Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;

Qy 5 VLLLTALTLCHFNLDENAMTFQENARGFQSVQLOGSRVVGVAPOBIVAAVNRGSLY 64
 Db 5 VLLLSVLAHYGFNLNDEBPTTFQEDAGGFGQSVQFQGSRLVVGAPLEVVAANTGRLY 64
 Qy 65 QCDYSTGSCGPIRLQVPVAVNMSLGLSLAATSPQLLIACGPTVHQTCSNTYKGLCF 124
 Db 65 DCAAAATGMCQPIPLHTRPEAVNMSLGLTLAASTNGSRLIACGPTLHRVCGENSYKGSCL 124

Qy 125 LFGSNLRQOPKFPPEALRGCPQEDSDIAFLIDSGSIIIPHFRFMKEFYSTVMEQLKSK 184
 Db 125 LLGSRW-EIIQTVPDATPECPHOEMDI VFLIDSGSIDQNDNFQMKGFQVAVMGQEGTD 183
 Qy 185 TLFSLMOYSEEFRIHFTFEFQNNPNRSLVKPITOLLGRTHTATGIRKVVRELFINITNG 244
 Db 184 TLFALMOYSNLLKHIFHTFTQFRTSPSQSLVDPIVLQGLTFTATGILTVVLTQLFHKNG 243
 Qy 245 ARKNAFKILVITDGKFGDPLGYEDVIEADREGVIRYVIGVDGAFRSEKSRQELNTIA 304
 Db 244 ARKSAKKILVITDGQKYKDPLEYSVDIPOAEKAGIIRIYAIGVGHAFQGTARQELNTIS 303
 Qy 305 SKPEPRHVFQNNFEALKTIQNLQREKIFAIEGTQTQTCSSSFEHEMSQEGFSAAITNGP 364
 Db 304 SAPQDHPVKVDNFALGSIQKQLEKIYAVEGTQSRASSSPQHENSQEGFSTALTMGDL 363
 Qy 365 LLSTVGSYDWAGVFLYTSKEKSTFINMTRVDSMDMDAYLCYAAAIIILNRVOSLVLAGP 424
 Db 364 FLGAVGSFSGGAFLYPPNMSPTFINMSQENYMDRDSYLGSTELALMKGVQNLVLGAP 423
 Qy 425 RYOHIGLVANFRONTGMWESNANVKTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYE 484
 Db 424 RYOHIGKAVITQVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLGAPHYE 483
 Qy 485 QTRGGQSVGCPPLPRGQARMCDAVLYGEOQPMGRFGAALTVLGVDVNGDKLTDAVITGAP 544
 Db 484 QTRGGQSVGCPPLPRGQARMCDAVLYGEOQPMGRFGAALTVLGVDVNGDKLTDAVITGAP 543
 Qy 545 GEDNRAVLYFHTSGSGISPSHSQRIAGSKLSPLQYFGOSLSSGGQDLTMDGLDVLTV 604
 Db 544 GEQENRAVLYFHTSGSGISPSHSQRIAGSKLSPLQYFGOSLSSGGQDLTMDGLDVLTV 603
 Qy 605 GAQGHVLLRSQPLRVKAIWENFNEPREVARNVPCNDQVVKGEAGVRVCHLVOKSTRD 664
 Db 604 GARGQVLLRSQPLRVKGVAMRFPVEVAKAVTRCWEKESALEAGQATVCLTITQKSSLD 663
 Qy 665 RLREGQIQSVVTVYDLALDSGRPHSRVAFNETQNSTRRQTQVGLHTOTCETLKLQLPNCIE 724
 Db 664 QL--GDIQSSVRFDLALDPGLRLTSRAIFNETQNTPLTRRKTGLGHCETLKLQLPDCE 721
 Qy 725 DPVSPVLRNLSVGLVPLSAFNGNLRPVLAEDAQRFTALFPFPEKNCNDNI CODDLSIT 784
 Db 722 DWSPPIILHLNLSLVREPISQNLRLPVLA VGSQDLFTASLPFPEKNCQDGLCEGLGVT 781
 Qy 785 FSNLSCLVVGVPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRORSW 844
 Db 782 LSFSGLOTLTVGSSLELNVIITVMNAGEDSVTVVSLYPAGLSHRRVSGNQKOPHQSAL 841
 Qy 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
 Db 842 RLACETV-PTDEG-LRSSRCSVNHPIFHEGSGNGTFIVTDFVSYKATLGDRLMRASASS 899
 Qy 905 ENNMPRTNKTFFOLELPVKYAVYVMVTSYHGVSTKYLPF-TASENTRVMQHOYQVSNLQ 963
 Db 900 ENNKASSSKATFOLELPVKYAVYTMISROEESTKYFNATSDKKMKEAEHRYVNNLSQ 959
 Qy 964 RSLPISLVLPVRLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVC 1023
 Db 960 RDLAISINFWPVLNGVAVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDC 1017
 Qy 1024 STIAVCORIQCDDIPFGIQQEENATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLL 1083
 Db 1018 SIADCLQPRCDVPSPSQEELDFTLKGNSLPGWVRETLQKKVLVVSAEITFTDTSVSQL 1077
 Qy 1084 PQGAFVRSOTETKVEPEVNPPLIVGSSVGLLALLALITAAIYKLGFFKROYKDMMS 1143
 Db 1078 PQEAFMRAQMEMVLEEDVYNAIPIIMGSSVGALLLALLITATLYKLGFFKRYKEMLE 1137
 Qy 1144 E 1144
 Db 1138 D 1138

Tue Nov 25 14:51:51 2003

us-09-902-481b-1.rapb

Page 13

Search completed: November 25, 2003, 14:41:06
Job time: 33.4941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:13:09 ; Search time 13.7527 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956
Sequence: 1 MALRVLLLTALTLCGHNLD.....FKROYKDMSEGGPGAPBPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5956	100.0	1153	1	US-08-173-497-3
2	5956	100.0	1153	1	US-08-286-889-3
3	5956	100.0	1153	1	US-08-485-618-3
4	5956	100.0	1153	1	US-08-362-652-3
5	5956	100.0	1153	2	US-08-605-672-3
6	5956	100.0	1153	2	US-08-482-293A-3
7	5956	100.0	1153	2	US-08-943-363-3
8	5956	100.0	1153	3	US-09-193-043-3
9	5956	100.0	1153	4	US-09-688-307A-3
10	5925.5	99.5	1152	2	US-08-476-062A-43
11	5925.5	99.5	1152	5	PCT-US96-01314-43
12	5925.5	99.5	1152	6	5424399-2
13	3503	58.8	1163	2	US-08-476-062A-44
14	3503	58.8	1163	5	PCT-US96-01314-44
15	3480	58.4	1163	1	US-08-173-497-4
16	3480	58.4	1163	1	US-08-286-889-4
17	3480	58.4	1163	1	US-08-485-618-4
18	3480	58.4	1163	1	US-08-362-652-4
19	3480	58.4	1163	2	US-08-605-672-4
20	3480	58.4	1163	2	US-08-482-293A-4
21	3480	58.4	1163	2	US-08-943-363-4
22	3480	58.4	1163	3	US-09-193-043-4
23	3480	58.4	1163	4	US-09-688-307A-4
24	3455	58.0	1161	1	US-08-173-497-2
25	3455	58.0	1161	1	US-08-286-889-2
26	3455	58.0	1161	1	US-08-485-618-2
27	3455	58.0	1161	1	US-08-362-652-2

28 3455 58.0 1161 2 US-08-605-672-2 Sequence 2, Appli
29 3455 58.0 1161 2 US-08-482-293A-2 Sequence 2, Appli
30 3455 58.0 1161 2 US-08-943-363-2 Sequence 2, Appli
31 3455 58.0 1161 3 US-09-193-043-2 Sequence 2, Appli
32 3455 58.0 1161 4 US-09-688-307A-2 Sequence 2, Appli
33 3439.5 57.7 1161 1 US-08-485-618-99 Sequence 99, Appli
34 3439.5 57.7 1161 2 US-08-605-672-99 Sequence 99, Appli
35 3439.5 57.7 1161 2 US-08-482-293A-99 Sequence 99, Appli
36 3439.5 57.7 1161 2 US-08-943-363-99 Sequence 99, Appli
37 3439.5 57.7 1161 3 US-09-193-043-99 Sequence 99, Appli
38 3439.5 57.7 1161 4 US-09-688-307A-99 Sequence 99, Appli
39 3270.5 54.9 1161 3 US-09-193-043-55 Sequence 55, Appli
40 3270.5 54.9 1161 4 US-09-688-307A-55 Sequence 55, Appli
41 3264 54.8 1161 1 US-08-485-618-53 Sequence 53, Appli
42 3264 54.8 1161 1 US-08-362-652-53 Sequence 53, Appli
43 3264 54.8 1161 2 US-08-605-672-53 Sequence 53, Appli
44 3264 54.8 1161 2 US-08-482-293A-53 Sequence 53, Appli
45 3264 54.8 1161 2 US-08-943-363-53 Sequence 53, Appli

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/173.497
; APPLICATION NUMBER: US/08/173.497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFQGSVVQLGSRVVVGAPOEIVAAHQ 60

Db 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFQGSVVQLGSRVVVGAPOEIVAAHQ 60

QY 61 GSYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
DB 61 GSYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCLFGLSGLNRQOQPKQPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
DB 121 GLCLFGLSGLNRQOQPKQPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
QY 181 KSKTFLSMLQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KSKTFLSMLQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQEL 300
DB 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIONQIREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
DB 301 NTIASKPRDHVQVNNFEALKTIONQIREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPLSTVGYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420
DB 361 SNGPLSTVGYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420
QY 421 LGAPRYOHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLGAP 480
DB 421 LGAPRYOHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLGAP 480
QY 481 HYYEQTGRGQVSVCPLRGRQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVA 540
DB 481 HYYEQTGRGQVSVCPLRGRQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVA 540
QY 541 IGARCEENRGAIVLPHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLV 600
DB 541 IGARCEENRGAIVLPHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLV 600
QY 601 DLTGAQGHVLLLSQOVLRYKATMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQK 660
DB 601 DLTGAQGHVLLLSQOVLRYKATMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQK 660
QY 661 STRDLRREGQIQSVVYTDALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLP 720
DB 661 STRDLRREGQIQSVVYTDALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPVILRNFLSVGTPLSAFNLRLPVLAEDAORLFTALPPEKNCNNDNICDD 780
DB 721 NCIEDPVSPVILRNFLSVGTPLSAFNLRLPVLAEDAORLFTALPPEKNCNNDNICDD 780
QY 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQORS 840
DB 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQORS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNPRNTKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQHQYQVSN 960
DB 901 NVTSENNPRNTKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQHQYQVSN 960
QY 961 LGORSLPLSLVFLVPVRLNQTVMIDRPOVTSSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGORSLPLSLVFLVPVRLNQTVMIDRPOVTSSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCQRIQCDIPFGIOEEFNATLKGSLSDWYIKTSHNHLIIIVSTAEIIFNDSVF 1080
DB 1021 VNCSTAVCQRIQCDIPFGIOEEFNATLKGSLSDWYIKTSHNHLIIIVSTAEIIFNDSVF 1080
QY 1081 TLLPQCGAFVRSQTTKTEPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQRYKD 1140
DB 1081 TLLPQCGAFVRSQTTKTEPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQRYKD 1140

QY 1141 MMSEGPPGAEPO 1153
DB 1141 MMSEGPPGAEPO 1153
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-286-889-3
Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLLLTALTLCCHFNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQBEIVAAQR 60
DB 1 MALRVLLLTALTLCCHFNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQBEIVAAQR 60
QY 61 GSYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
DB 61 GSYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCLFGLSGLNRQOQPKQPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
DB 121 GLCLFGLSGLNRQOQPKQPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
QY 181 KSKTFLSMLQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KSKTFLSMLQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQEL 300
DB 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQEL 300

QY 301 NTIASKPPRDHVPQVNNFEALKTIONQREKIFAIBGTQTGSSSSFEHMSQSGFSAIT 360
DB 301 NTIASKPPRDHVPQVNNFEALKTIONQREKIFAIBGTQTGSSSSFEHMSQSGFSAIT 360
QY 361 SNGPLLSYVGSYDAGVFLYTSKSKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
DB 361 SNGPLLSYVGSYDAGVFLYTSKSKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMFQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYIEQTRGGQVSVCPILPRQARAWQCDVLYGQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYIEQTRGGQVSVCPILPRQARAWQCDVLYGQGPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEDNRCAYVLFHGTSGSISPSHSORIASGLSPRLQYFGQSLGGQDLTWDGLV 600
DB 541 IGAPGEDNRCAYVLFHGTSGSISPSHSORIASGLSPRLQYFGQSLGGQDLTWDGLV 600
QY 601 DLTVGAQGHVLLRSOPVLKVAIMEFNPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
DB 601 DLTVGAQGHVLLRSOPVLKVAIMEFNPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
QY 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSLCTPLSAGNLRPLVLAEDAQRFTALFPFKNCNGNDICQDD 780
DB 721 NCIEDPVSPIVLRNLSLCTPLSAGNLRPLVLAEDAQRFTALFPFKNCNGNDICQDD 780
QY 781 LSITFSFMSLDCLVVGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
DB 781 LSITFSFMSLDCLVVGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPTNTEFOLELPVKYAVYVWVTSYGVSTKYLNFTASENTSRVMOHQVSN 960
DB 901 NVTSENNMPTNTEFOLELPVKYAVYVWVTSYGVSTKYLNFTASENTSRVMOHQVSN 960
QY 961 LGORSPLISLVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
DB 961 LGORSPLISLVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
QY 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSF 1080
DB 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSF 1080
QY 1081 TLLPGQGAFAVRSOTETKVEFVFNPLPLIVGSSVGLLALLITAAIYKLGFFKQYKD 1140
DB 1081 TLLPGQGAFAVRSOTETKVEFVFNPLPLIVGSSVGLLALLITAAIYKLGFFKQYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153

RESULT 3

US-08-485-618-3

; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vliet, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANOR 60
DB 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANOR 60
QY 61 GSYLQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
DB 61 GSYLQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLPGSNLRQOPQKFPPEALRCQPEDSDIAFLIDGSGSIIPHDPRMKKEFYSTVMEQL 180
DB 121 GLCFLPGSNLRQOPQKFPPEALRCQPEDSDIAFLIDGSGSIIPHDPRMKKEFYSTVMEQL 180
QY 181 KSKTLFSLMOYSEEFRIHFTFKEFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
DB 181 KSKTLFSLMOYSEEFRIHFTFKEFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVPQVNNFEALKTIONQREKIFAIBGTQTGSSSSFEHMSQSGFSAIT 360
DB 301 NTIASKPPRDHVPQVNNFEALKTIONQREKIFAIBGTQTGSSSSFEHMSQSGFSAIT 360
QY 361 SNGPLLSYVGSYDAGVFLYTSKSKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
DB 361 SNGPLLSYVGSYDAGVFLYTSKSKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480

```
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Qy 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTGAGQGHVLLRSQPLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660
Db 601 DLTGAGQGHVLLRSQPLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660
Qy 661 STRDLREGQIQSVVTYDIALDSGRPHSAFVNETKNSRRTQVGLTQTCTETLKLQLP 720
Db 661 STRDLREGQIQSVVTYDIALDSGRPHSAFVNETKNSRRTQVGLTQTCTETLKLQLP 720
Qy 721 NCIEDPVSPIVLRNPLSVGTPLSAFNGLRPVLAEDAQLFTALPPFEKNCNDNICODD 780
Db 721 NCIEDPVSPIVLRNPLSVGTPLSAFNGLRPVLAEDAQLFTALPPFEKNCNDNICODD 780
Qy 781 LSTTFMSLDCLVCGPREFNVTVVRNDGDSYRTQVTFEPLDLSYRKYSTLONORS 840
Db 781 LSTTFMSLDCLVCGPREFNVTVVRNDGDSYRTQVTFEPLDLSYRKYSTLONORS 840
Qy 841 QRSWRLACASASTSVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACASASTSVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNMPTNKTEFQLELPVKYAVYVMVTSKYLNTFASNTSRVNVQHQYQVSN 960
Db 901 NVTSENNMPTNKTEFQLELPVKYAVYVMVTSKYLNTFASNTSRVNVQHQYQVSN 960
Qy 961 LQORSPLISLVFLVPLRNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LQORSPLISLVFLVPLRNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCISAVCORIOCDIPFGIQEENATLKNLSFDWYIKTSHNHLIIVSTAILFNDSVF 1080
Db 1021 VNCISAVCORIOCDIPFGIQEENATLKNLSFDWYIKTSHNHLIIVSTAILFNDSVF 1080
Qy 1081 TLLPGGCAFVRSQTEKVPFEPVNPPLPLIVSGSVGLLLALITAALYKLGFFKRYQKD 1140
Db 1081 TLLPGGCAFVRSQTEKVPFEPVNPPLPLIVSGSVGLLLALITAALYKLGFFKRYQKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153
```

RESULT 4

```
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362.652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286.889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: William Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCGHNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQR 60
Db 1 MALRVLLLTALTLCGHNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQR 60
Qy 61 GSLVQCDYSTGSCSPIRLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSENTVVK 120
Db 61 GSLVQCDYSTGSCSPIRLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSENTVVK 120
Qy 121 GLCLFGLGNLRQQQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQL 180
Db 121 GLCLFGLGNLRQQQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQL 180
Qy 181 KKSCTLPSLMQYSSEFRIFHTFKFQNNPNRSLVKPTQLLGRTHATGIRKVVRELFPN 240
Db 181 KKSCTLPSLMQYSSEFRIFHTFKFQNNPNRSLVKPTQLLGRTHATGIRKVVRELFPN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
Qy 361 SNGPLLSTVSGYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLV 420
Db 361 SNGPLLSTVSGYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Qy 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTGAGQGHVLLRSQPLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660
Db 601 DLTGAGQGHVLLRSQPLRVKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660
```

601 DLTGAGHVVLLRSQVLRVKAIMEFNPENPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTETLKQLP 720
661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTETLKQLP 720
721 NCIEDPVPSPVLRNLSLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
721 NCIEDPVPSPVLRNLSLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
781 LSITFSFMSLDCVLVGGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840
781 LSITFSFMSLDCVLVGGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840
841 QRSWLACASASSTEVSGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKA 900
841 QRSWLACASASSTEVSGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKA 900
901 NVTSENMPRTNKTETQLELPLVKYAVYVMTSHGVSTKYLNFTASNTSRVQHQVQSN 960
901 NVTSENMPRTNKTETQLELPLVKYAVYVMTSHGVSTKYLNFTASNTSRVQHQVQSN 960
961 LGORSPLISLVLPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
961 LGORSPLISLVLPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
1021 VNCISIAVCQRIQCDIPFFQIEBFNATLKGNSLFDWYIKTSHNHLIVSTAELLFNDVSF 1080
1021 VNCISIAVCQRIQCDIPFFQIEBFNATLKGNSLFDWYIKTSHNHLIVSTAELLFNDVSF 1080
1081 TLLPGGAFVRSOTETKVPFEPVNPPLPIVGVSGGLLALLITAAALKYKGFKKQYKD 1140
1081 TLLPGGAFVRSOTETKVPFEPVNPPLPIVGVSGGLLALLITAAALKYKGFKKQYKD 1140
1141 MMSEGGPPGAEPQ 1153
1141 MMSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
Sequence 3, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-3
Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANOR 60
DB 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANOR 60
QY 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSSENTYVK 120
DB 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSSENTYVK 120
QY 121 GLCFIFGSLNRQOPKQPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOL 180
DB 121 GLCFIFGSLNRQOPKQPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOL 180
QY 181 KSKTLFSLMOYSEBFHFTFKFQNNPNRSLVKPITQLGRTHTATGIRKVVRELPN 240
DB 181 KSKTLFSLMOYSEBFHFTFKFQNNPNRSLVKPITQLGRTHTATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
DB 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 360
QY 361 SNGPLLSVTVGSYDMAGGVFLYTSKSTFINMTRVDSMDNDAYLVGAAILLRNRVQSLV 420
DB 361 SNGPLLSVTVGSYDMAGGVFLYTSKSTFINMTRVDSMDNDAYLVGAAILLRNRVQSLV 420
QY 421 LGAPRYQHIGLVAMPRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMPRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEOTRGQSVSCPLPRGQARWQCDVLYGEQSQPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVSCPLPRGQARWQCDVLYGEQSQPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 ICAPGEENRGAVLPHGTSGSGISPSHSQRTAGSKLSPRLOYFGQSLGGODLTMDGLV 600
DB 541 ICAPGEENRGAVLPHGTSGSGISPSHSQRTAGSKLSPRLOYFGQSLGGODLTMDGLV 600
QY 601 DLTGAQGHVLLLRSQPVLRVKAIMEFNPENPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
DB 601 DLTGAQGHVLLLRSQPVLRVKAIMEFNPENPREVARNVFECNDQVVKKEAGEVRVCLHVOK 660
QY 661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTETLKQLP 720
DB 661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCTETLKQLP 720
QY 721 NCIEDPVPSPVLRNLSLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
DB 721 NCIEDPVPSPVLRNLSLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780

Qy 781 LSITFSMSLDCLVVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSMSLDCLVVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRTNKTFFOLELPKVAVYVMVTSHGVSSTKYLNFTASENTSRVMOHOYQVSN 960
Db 901 NVTSENNPRTNKTFFOLELPKVAVYVMVTSHGVSSTKYLNFTASENTSRVMOHOYQVSN 960
Qy 961 LQGRSLPISLVFLVPRLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPV 1020
Db 961 LQGRSLPISLVFLVPRLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPV 1020
Qy 1021 VNCSTAVCQRIQCDDIPFGIOEFENATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSF 1080
Db 1021 VNCSTAVCQRIQCDDIPFGIOEFENATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSF 1080
Qy 1081 TLLPQGGAFVRSQETKVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKROYKD 1140
Db 1081 TLLPQGGAFVRSQETKVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKROYKD 1140
Qy 1141 MMSEGGPPGABPQ 1153
Db 1141 MMSEGGPPGABPQ 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 381659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-3
Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTICGHNLDTENAMTFQENARFGQSVVQLQCSRVVVVGAPQEIIVANQR 60
Db 1 MALRVLLLTALTICGHNLDTENAMTFQENARFGQSVVQLQCSRVVVVGAPQEIIVANQR 60
Qy 61 GSIYQCDYSYSGCEPRLQVPVEAVNMSLGLSLAATPPQLLACGPTVHQTCSNTYVK 120
Db 61 GSIYQCDYSYSGCEPRLQVPVEAVNMSLGLSLAATPPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFLFSGNLRRQPKPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSTVMEQL 180
Db 121 GLCFLFSGNLRRQPKPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSTVMEQL 180
Qy 181 KSKTFLPSLMQYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHTATGIRKVVRELFN 240
Db 181 KSKTFLPSLMQYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHTATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFOVNNFEALKTIONQLREKIIPAIECTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPPRDHVFOVNNFEALKTIONQLREKIIPAIECTQTGSSSSFEHMSQEGFSAIT 360
Qy 361 SNGPLLTSGSYDWAGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLTSGSYDWAGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy 421 LGAPRYOHIGLVAMFRONTGWNESNANVKGTOICAYFGASLCSVDVDSNGSTDVLICAP 480
Db 421 LGAPRYOHIGLVAMFRONTGWNESNANVKGTOICAYFGASLCSVDVDSNGSTDVLICAP 480
Qy 481 HYYEQTRGGQVSVCPPLRGQARQWQCDALVYGEQGPWGRFGAALTVLGDNVNGDKLTDVA 540
Db 481 HYYEQTRGGQVSVCPPLRGQARQWQCDALVYGEQGPWGRFGAALTVLGDNVNGDKLTDVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLMGDLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLMGDLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
Qy 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKSTRQTQVGLGTQTCETLKLQLP 720
Db 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKSTRQTQVGLGTQTCETLKLQLP 720
Qy 721 NCIEDPVSPVLRNLNFSLVGTPLSAFNLRPVLAEDAQRLETFALPFPEKNCNDNICDD 780
Db 721 NCIEDPVSPVLRNLNFSLVGTPLSAFNLRPVLAEDAQRLETFALPFPEKNCNDNICDD 780
Qy 781 LSITFSMSLDCLVVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSMSLDCLVVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRTNKTFFOLELPKVAVYVMVTSHGVSSTKYLNFTASENTSRVMOHOYQVSN 960

Db 901 NVTSENNMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYLNFASNTSRVMQHQYQVSN 960
Qy 961 LGQORSLPISLVFLVPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
Db 961 LGQORSLPISLVFLVPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDVSF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDVSF 1080
Qy 1081 TLLPGGAFVRSOTETKVPFEPVNPPLIVGVSSVGLLILLALITAAVLKLGFFKQYKD 1140
Db 1081 TLLPGGAFVRSOTETKVPFEPVNPPLIVGVSSVGLLILLALITAAVLKLGFFKQYKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-943-363-3

Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAQR 60
Db 1 MALRVLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAQR 60
Qy 61 GSLYOCDSYSGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVK 120
Db 61 GSLYOCDSYSGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVK 120
Qy 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Qy 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Db 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAPRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAPRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRNVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRNVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYTEOTRGQSVSVCPPLRGQARWOCDAVLVGEQCPQWGRFCAALTVLGDVNGDKLTQVA 540
Db 481 HYTEOTRGQSVSVCPPLRGQARWOCDAVLVGEQCPQWGRFCAALTVLGDVNGDKLTQVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDNDVVVKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDNDVVVKEAGEVRVCLHVOK 660
Qy 661 STRDRLRGQIQSVVTVYDLALDSGRPHSRVFNENKSTRROTQVGLTQTCETLKLQLP 720
Db 661 STRDRLRGQIQSVVTVYDLALDSGRPHSRVFNENKSTRROTQVGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPIVLRLNFSLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFKNCNDNICQDD 780
Db 721 NCIEDPVSPIVLRLNFSLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFKNCNDNICQDD 780
Qy 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNR 840
Db 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNR 840
Qy 841 QRSWELACESASSTEVSGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWELACESASSTEVSGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYLNFASNTSRVMQHQYQVSN 960
Db 901 NVTSENNMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYLNFASNTSRVMQHQYQVSN 960
Qy 961 LGQORSLPISLVFLVPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
Db 961 LGQORSLPISLVFLVPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDVSF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDVSF 1080

SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 100.0%; Score 5956; DB 4; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLGSRVVVGAPOEIVANQR 60
DB 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLGSRVVVGAPOEIVANQR 60
QY 61 GSLYQCDYGTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYGTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLFGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
DB 121 GLCFLFGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
QY 181 KSKTLFSLMQYSEEPRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
DB 181 KSKTLFSLMQYSEEPRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYGVGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYGVGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVPQNNFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAIT 360
DB 301 NTIASKPPRDHVPQNNFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAIT 360
QY 361 SNGPLSTVGSYDMAGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSIV 420
DB 361 SNGPLSTVGSYDMAGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSIV 420
QY 421 LGAPRYQHLVAMPQONTGMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
DB 421 LGAPRYQHLVAMPQONTGMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEOTRGQSVSCPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVSCPLPRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVYLPFGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV 600
DB 541 IGAPGEDNRGAVYLPFGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLV 600
QY 601 DLTVGAGQHVLLLRSPVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTVGAGQHVLLLRSPVLRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGIOISVVTYDLALDSGRPHSAVFNETKNSTRROTQVLGLTQTCETLKLQLP 720
DB 661 STRDLREGIOISVVTYDLALDSGRPHSAVFNETKNSTRROTQVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSLVGTPLSAFNGLRPVLAEADAQRLLFTALPFPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNLSLVGTPLSAFNGLRPVLAEADAQRLLFTALPFPEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVGQPREFNTVTVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRS 840
DB 781 LSITFSFMSLDCLVGQPREFNTVTVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRS 840
QY 841 QRSWLACASASTEVSGALKSTCSINHPIPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACASASTEVSGALKSTCSINHPIPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFQLELPVKYAVYVMVTVSHGVSTKYLNFTASNTSRVVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTEFQLELPVKYAVYVMVTVSHGVSTKYLNFTASNTSRVVMQHOYQVSN 960
```

```
QY 961 LQORSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPV 1020
DB 961 LQORSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPV 1020
QY 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1080
DB 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1080
QY 1081 TLLPGQGAFAVRSQETKVEPEFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKD 1140
DB 1081 TLLPGQGAFAVRSQETKVEPEFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153
```

RESULT 10

```
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-476-062A-43
```

Query Match 99.5%; Score 5925.5; DB 2; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLGSRVVVGAPOEIVANQR 60

Db 1 MALRVLLLTALTLCHGNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVANQR 60
QY 61 GSYLQCDYSTGSCPTRLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
Db 61 GSYLQCDYSTGSCPTRLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCFLFGSNLRQOQPKFPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOQPKFPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
QY 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
Db 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQBGFSAAIT 360
Db 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQBGFSAAIT 360
QY 361 SNGPLLSTVGSYDMAGGVFLYTSKESKSTFINMTRVSDMDNDAYLGYAAAIILNRVQSLV 420
Db 361 SNGPLLSTVGSYDMAGGVFLYTSKESKSTFINMTRVSDMDNDAYLGYAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
QY 481 HYIEOTRGGOVSVCLPRGORARWQCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVA 540
Db 481 HYIEOTRGGOVSVCLPRG-RARWQCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVA 539
QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYFGQSLGGQDLTMDGLV 600
Db 540 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYFGQSLGGQDLTMDGLV 599
QY 601 DLTVGAQGHVLLRSQVLRVKAIMEFNPREVARNFECDQVVKGEAGEVRCVCLHVOK 660
Db 600 DLTVGAQGHVLLRSQVLRVKAIMEFNPREVARNFECDQVVKGEAGEVRCVCLHVOK 659
QY 661 STRDLREGOIQSVVTVYDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLP 720
Db 660 STRDLREGOIQSVVTVYDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLP 719
QY 721 NCIEDPVSPIVLRNLFSLVGTPLSAFCNLRPVLAEDAQRFTALFPPEKNCNDNICQDD 780
Db 720 NCIEDPVSPIVLRNLFSLVGTPLSAFCNLRPVLAEDAQRFTALFPPEKNCNDNICQDD 779
QY 781 LSITFSPMSLDCLVVGPREBNVTVYRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRS 840
Db 780 LSITFSPMSLDCLVVGPREBNVTVYRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRS 839
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 840 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 899
QY 901 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
Db 900 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 959
QY 961 LGQRSPLISLVLPVPRNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
Db 960 LGQRSPLISLVLPVPRNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1019
QY 1021 VNCISIAVCQRIQCQDIPFEGIOEFENATLKGNSLFDWIKTSHNHLIVSTAEIIFNDSVF 1080
Db 1020 VNCISIAVCQRIQCQDIPFEGIOEFENATLKGNSLFDWIKTSHNHLIVSTAEIIFNDSVF 1079
QY 1081 TLLPQOGAFVRSQSTETKVEPPEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYOKD 1140

Db 1080 TLLPQOGAFVRSQSTETKVEPPEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYOKD 1139
QY 1141 MMSEGGPPGABPO 1153
Db 1140 MMSEGGPPGABPO 1152
RESULT 11
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnsout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US96-01314-43

Query Match 99.5%; Score 5925.5; DB 5; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MALRVLLLTALTLCHGNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVANQR 60
Db 1 MALRVLLLTALTLCHGNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVANQR 60
QY 61 GSYLQCDYSTGSCPTRLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
Db 61 GSYLQCDYSTGSCPTRLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCFLFGSNLRQOQPKFPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOQPKFPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
QY 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
Db 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300

301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFINTRVDSMDNDAYLGAAAIILNRVQSIV 420
361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFINTRVDSMDNDAYLGAAAIILNRVQSIV 420
421 LGAPRYQHIGLVAMFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
421 LGAPRYQHIGLVAMFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
481 HYYEOTRGQVSVCPPLRGORARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTQV 540
481 HYYEOTRGQVSVCPPLRG - RARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTQV 539
541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
540 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 599
601 DLTGVAQGHVLLRSQVPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660
600 DLTGVAQGHVLLRSQVPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 659
661 STRDLREGQIQSVVYTDALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQPL 720
660 STRDLREGQIQSVVYTDALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQPL 719
721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAORLFTALPFPEKXKNCNDNICQDD 780
720 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAORLFTALPFPEKXKNCNDNICQDD 779
781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGDSYRTQVTFEPFLDLSYRKVSTIQNQRS 840
780 LSITFSFMSLDCLVVGGPREFNVTVVRNDGDSYRTQVTFEPFLDLSYRKVSTIQNQRS 839
841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKA 900
840 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKA 899
901 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
900 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 959
961 LQORSUPISLVFLVPLVRLNQTIVDRPQVTFESNLSSTCTKTERLPSHDSFLAELRKA 1020
960 LQORSUPISLVFLVPLVRLNQTIVDRPQVTFESNLSSTCTKTERLPSHDSFLAELRKA 1019
1021 VNCISIAVCORIQCIDIFFPGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFND 1080
1020 VNCISIAVCORIQCIDIFFPGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFND 1079
1081 TLLPGOGAFVRSQTEKVPFVNPPLPIVGSVGGLLLLALITAAALKYLGFFKQYKX 1140
1080 TLLPGOGAFVRSQTEKVPFVNPPLPIVGSVGGLLLLALITAAALKYLGFFKQYKX 1139
1141 MMSEGGPPGAEPQ 1153
1140 MMSEGGPPGAEPQ 1152

RESULT 12

5424399-2

; Patent No. 5424399

; APPLICANT: ARNAOUT, M. AMIN

; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS

; NUMBER OF SEQUENCES: 12

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/78,871

; FILING DATE: 16-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 539,842

; FILING DATE: 18-JUN-1990

; APPLICATION NUMBER: 212,573

; FILING DATE: 28-JUN-1988

; SEQ ID NO: 2:

; LENGTH: 1152

5424399-2

Query Match 99.5%; Score 5925.5; DB 6; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MALRVLLLTALTLCCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI VAAANOR 60
DB 1 MALRVLLLTALTLCCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI VAAANOR 60
QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHOTCSENTYVK 120
DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHOTCSENTYVK 120
QY 121 GLCFUFGNLRQOPQKFPPEARLRCGQEDSDIAFLIDGSGSIIPHDFRMRKEFPVSTMEO 180
DB 121 GLCFUFGNLRQOPQKFPPEARLRCGQEDSDIAFLIDGSGSIIPHDFRMRKEFPVSTMEO 180
QY 181 KKSKTLPFSLMOYSEBEFRIHPTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELPN 240
DB 181 KKSKTLPFSLMOYSEBEFRIHPTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
DB 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFINTRVDSMDNDAYLGAAAIILNRVQSIV 420
DB 361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFINTRVDSMDNDAYLGAAAIILNRVQSIV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEOTRGQVSVCPPLRGORARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTQV 540
DB 481 HYYEOTRGQVSVCPPLRG - RARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTQV 539
QY 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
DB 540 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 599
QY 601 DLTGVAQGHVLLRSQVPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660
DB 600 DLTGVAQGHVLLRSQVPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 659
QY 661 STRDLREGQIQSVVYTDALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQPL 720
DB 660 STRDLREGQIQSVVYTDALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQPL 719
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAORLFTALPFPEKXKNCNDNICQDD 780
DB 720 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAORLFTALPFPEKXKNCNDNICQDD 779
QY 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGDSYRTQVTFEPFLDLSYRKVSTIQNQRS 840
DB 780 LSITFSFMSLDCLVVGGPREFNVTVVRNDGDSYRTQVTFEPFLDLSYRKVSTIQNQRS 839
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKA 900
DB 840 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKA 899
QY 901 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
DB 900 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 959

Qy 1144 E 1144
Db 1143 E 1143

RESULT 14

PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-44

Query Match 58.8%; Score 3503; DB 5; Length 1163;
Best Local Similarity 61.3%; Pred. No. 4.2e-287;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;
Qy 5 VLLLTALTLCHGNLDTEENAMTQENARGQSVVQLQSGSRVVGAPQEIIVANQSGSLY 64
Db 8 LLLLTALTSGLGNLDTEELTAFRVDSAGGDSVVQYANVWVVGAPQKITAANQGTGLY 67
Qy 65 QCYSTGCEPIRLQVPEAVNNSGLSLAATSPPOLACGPTVHGTCSNTYVXGLCF 124
Db 68 QCYSTGACBPGLQVPEAVNNSGLSLASTTSPSOLLACGPTVHHECGRNMYLTGLCF 127
Qy 125 LFGSNLRQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKS 184
Db 128 LLGPT--QLTORLPVSRQECPRQEQDIVFLIDGSGSISSRNFMATMNFVRAVISQFORPS 185
Qy 185 TLFSLMQYSEFRIHFTFKFQNNPRSLVKPITQLLGHATGIRKVVRELFINITNG 244
Db 186 TQSLMQFSNKFQTHFTFEERFSTNPLSLASVHQLQGYTATATQNVVHRLPHASYG 245
Qy 245 ARNAFKILVITDGKFGDPLGVEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIA 304
Db 246 ARDARKILVITDGKFGDPLGVEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIA 305
Qy 305 SKPPRDHVQNNFEALKTQNLQREKIFAIEGTQTGSSSSSFEHMSQEGFSAAITNGP 364

Db 306 SKPSOEHIKPKVDFDALKDIOQLKEKIFAIEGTQTTSSSSFELEMAQGFSAVFPDGP 365
Qy 365 LLSTVGSYDWAGGVFLYTSKESKSTFINMTRVDSDDMDAYLGAAAAIILNRVQSLVGLAP 424
Db 366 VLGAVGSFTMSGGAFLYPPNMSPTFINMSQENVDMEDSYLGSTELALWKGVQSLVGLAP 425
Qy 425 RYQHIGLVAMFRONTQWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIE 484
Db 426 RYQHTGKAVIFTQVSRQWRMKAETGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYIE 485
Qy 485 QTRGGQVSVCPPLRGORARWQCDVLYGQSGOPWGRFGAALTVLGDNVNGDKLTDVAIGAP 544
Db 486 QTRGGQVSVCPPLRGWR--RWWCDVLYGQSGOPWGRFGAALTVLGDNVNGDKLTDVVGAP 544
Qy 545 GEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTV 604
Db 545 GREENRGAVYLFHGVLGPSISPSHSQRIAGSKLSPRLQYFGQALSGQDLTQDGLVDLAV 604
Qy 605 GAQGHVLLRSQPLRVKAIMFNPREVARNVFECDNQVVKGEAGEVRLHVQKSTRD 664
Db 605 GARGQVLLLRTRPVLWVGVMQFIPAEIPRSAPFECREQVVEQTLVQSNICLYIDKRSKN 664
Qy 665 RLREGQIOSVTVYDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQNPCLIE 724
Db 665 LLGSRDLQSVTLDLALDPGRLSPRATFOETKNSRSLRVRLGKAKHCENFLLPSCVE 724
Qy 725 DPVSPITVRLNFSLVGTPLSAFQNLRPVLAEDAQRLLFTALFPFEKCNKGNINICQDLSIT 784
Db 725 DSVPTITLRLNFTLVGKPLAFRNLRPMLAALAAQRYFTASLFEKNCNGADHICQDNLGIS 784
Qy 785 FGFMSLDCLVVGCPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRW 844
Db 785 FGFPGKLSLLVGSNLELNAEVMVMNDGEDSYGTTITFSHPAGLSYRYVAEGOKQGLRSL 844
Qy 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTCDSPVGG--SQCTWSTSCRNHLIFRGGQITFLATFDVSPKAVLGDRLLLTANVS 902
Qy 905 ENNMPTNKTEFOLELPVKYAVVMVVTSHGVSTKYNLTAS--ENTSRVMOHOYQVSNLQ 963
Db 903 ENNTPTSTKTTFOLELPVKYAVYTVVSSHEQTKYLNFSSEKESHVAMHRVQVNNLQ 962
Qy 964 RSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPFSHSDFLAELRKAPVNC 1023
Db 963 RDLPSVSNFVWVPELVNQEAVMDEVSHPQNPSPSRCSSEKIAAPPASDFLAHIQKNPVLDC 1022
Qy 1024 STAVCORIOCDIPFGIOEENATLKGNSLSPDWYIKTSHNHLILVSTAILFNDSVFTLL 1083
Db 1023 SIAGCLRFCDVPFSVQBEELDFTLKGNLSFGVVRQILQKQVSVSVSAEITDTSVYSQL 1082
Qy 1084 PGQGAFFVRQOTETKVPFEPVNPPLPLIVGSSVGLLLALLAALVYKLGFFKQYKQDMMS 1143
Db 1083 PGQGAFFVRQOTETKVPFEPVNPPLPLIVGSSVGLLLALLAALVYKLGFFKQYKQDMMS 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 15

US-08-173-497-4
; Sequence 4, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60608-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,497
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 543795band, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31363
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-497-4

Query Match 58.4% Score 3480; DB 1; Length 1163;
Best Local Similarity 61.1% Pred. No. 3.7e-285;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

Qy	5	VLLTALTLCHGFLNLTENAMTPQENARGQSVVLOGSRVVGCAPOEIVAAQSGSLY	64
Db	8	LLLEFATATSLGFLNLTDELTAFRVDSAGFGDSVQVYANSWVVGAPQKIIAANQIGGLY	67
Qy	65	QCDYSTGCEPIRLQVPEAVNMISGLSLAATTSPQLACGPTVHOTCSENTVYKGLCF	124
Db	68	QCGYSTGACEPIGLQVPEAVNMISGLSLAATTSPQLACGPTVHCEGGRNWTGLCF	127
Qy	125	LFGSNLRQOPKFPALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKFVSTVMEQLKSK	184
Db	128	LLGPT--QLTQLPVSQCEPQEQDQIVFLIDGSGSISSRNPFATMNFVRAVISQFORPS	185
Qy	185	TLFSLMQYSEEPRIHFTKPEQNNPNRSLVKPITQLGRTHATGIRKRVRELPNITNG	244
Db	186	TOFSLMQFSNKFQTHFTFEERFRTSNPLSLASVHQLOQFTYTATAIQNVVHRLPHASYG	245
Qy	245	ARKNAFKILVITDGEKGDPLGEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIA	304
Db	246	ARRDAIKILVITDGGKGGSDLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKLNDIA	305
Qy	305	SKPPRDHVFQVNNFEALKTIONLREKIFAIBGTQTGSSSSPEHEMSQEGFSAATISNGP	364
Db	306	SKPSEHIFKVEDFDALQIQNLKEIFAIEGTETISSSSFELEWAQEGFSAVTDPGP	365
Qy	365	LLSTVGSYDAGGVFLYTSKESKTFINMTRVDSDDNDAYLGYAAAIILRNRVQSLVGLAP	424
Db	366	VLGAVGSFTWSGCAFLYPPNMPSTFINNSQENVMDRDSYLGYSTELALMKVGQSLVGLAP	425
Qy	425	RYQHIGLVAMPQNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYE	484
Db	426	RYOHIGKAVIFTQSVQRMRKAEVGTQIGSYFGASLCSVDVDTDGTDLVLIGAPHYYE	485
Qy	485	QTRGGQSVCPPLPRGORARWQCDVLYGEOGQWGRFGAALTVLGVDNGDKLTDVAIGAP	544
Db	486	QTRGGQSVCPPLPRGWR--RWCDVLYGEOGHPWGRFGAALTVLGVDNGDKLTDVVI GAP	544
Qy	545	GEEDNRGAVYLFPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLT	604
Db	545	GEENRGAVYLFPHGVLGPSISPSHSQRIAGSKLSSRLQYFGQALSGGQDLTQDGLVDLAV	604

Search completed: November 25, 2003, 14:23:09
Job time : 17.7527 secs

Qy	605	GAQGHVLLLRSPVLRVKAIMEFNPREVARNVECNQVVKVKGKAGEVRVCLHVQKSTRD	664
Db	605	GARGVLLLRTRPVLWVGVMQIPAEIPRSAFECEQVVSQTLVQSNICLYIDKRSKN	664
Qy	665	RLREGQIOSVVTYDLALDSGRPHSRVAFNETKNSRTQTVLGLTOTCETLKLQOLNCIE	724
Db	665	LLGSRDLQSSVTLDLALAPGRLSPRAFIOETKRSLSRVRLVGLKAHCENFNLLSPSCVE	724
Qy	725	DPVSPITVRLNLSLVTPISAFGLNRPVLAEDAQRIFTALFPPEKNCNDNICODLSIT	784
Db	725	DSVIPILRLNFTLVGKPLLAFLNRLPMLAALQRYFTASLPEKNCAGADHICODNLGIS	784
Qy	785	FSFMSLDCLVVGPRBFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQORSW	844
Db	785	FSFPGLKLLVGSNLELNAEVMVMDGEDSYGTITTFSHPAGLSYRYVAEGQKQGLRSL	844
Qy	845	RLACESASSTEVSGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTS	904
Db	845	HLTC--CSAPVGSQGTWSTSCRINHIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS	902
Qy	905	ENNMPTNKTEFOLELPVKYAVYVMTSHGVSTKYLNFTAS--ENTSRVMOHQYQVSNLQ	963
Db	903	ENNIPTSKTIFOLELPVKYAVYVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQ	962
Qy	964	RSIPISLVFLVRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNC	1023
Db	963	RDLPVSIWVWVVELNQEAVMMDVEVSHQPNSLRCSSEKIAPPASDLAHIQNPVLDC	1022
Qy	1024	SIACQRIQCDIPFFGLOEEFNATLKNLSFDMVYKTSNNHLLIVSTAELIENDSVFTLL	1083
Db	1023	SIAGCLFRCDVPFSFVQEEELDTLKNLSFGWVRQILQKKVSVSVVAEIIFDTSVYSQ	1082
Qy	1084	PGQAFVRSOTETKVEPFEVNPPLPLIVGSSVGLLLALITLITVLYKLGFPKRYKDMMS	1143
Db	1083	PGQAFVRSOTETKVEPFEVNPPLPLIVGSSVGLLLALITLITVLYKLGFPKRYKEMME	1142
Qy	1144	E 1144	
Db	1143	E 1143	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPFGEAPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying Chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum-Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5839	99.3	1153	1 RWHU1B	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.8	1163	1 RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 I56126	lymphocyte fuction
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1054	17.9	1181	2 A33998	integrin alpha-2 c
11	1049	17.8	1180	2 A35854	integrin alpha-1 c
12	666	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alaph integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44250	integrin alpha-5 c
20	551.5	9.4	1034	2 A36108	integrin alpha-v c
21	539	9.2	1044	2 T10050	integrin alpha-v c
22	533.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S16516	integrin alpha-8 c
29	497	8.5	1394	2 A29637	position-specific

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: Complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:g307148

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965,'P',967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA5

A:Note: The authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was con
R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A:Reference number: A32218; MUID:89098993; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FL>
A:Cross-references: GB:52227; NID:G263047; PIDN:AAB24821.1; PID:G263049
A:Note: the last three bases of intron 13, CAG, are included in some but not all mature
A:Note: sequence extracted from NCBI backbone (NCBIP:121963)
R:Pierce, M.W.; Remold-O'Donnell, B.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PI>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genet.Cs:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A:Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; ma
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.3% Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTQENARFGQSVVQLQSGRVVVGAPQEIWAANQRGLYQCDYSGSCEPI 60
Db 17 FNLDTENAMTQENARFGQSVVQLQSGRVVVGAPQEIWAANQRGLYQCDYSGSCEPI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNTGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNTGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQV 316

Qy 301 NFEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 376
Qy 361 GVFLYTSKSKSPFINNTRVDSMDNDAYLGAAAAIILRNVRQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSPFINNTRVDSMDNDAYLGAAAAIILRNVRQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPL 496
Qy 481 PRQQRARWQCDVLYGEQGPQWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 497 PRQQRARWQCDVLYGEQGPQWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVG 780
Db 737 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFEFPDLDSYRKYSTLONORSQSRWLACASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFEFPDLDSYRKYSTLONORSQSRWLACASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVWQHQVQVSNLGRORSIPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVWQHQVQVSNLGRORSIPISLVFLVPV 976
Qy 961 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISVACORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISVACORIQCDIP 1036
Qy 1021 FFGIOBEFNATLXGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLXGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096
Qy 1081 KVBPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGABPQ 1137
Db 1097 KVBPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGABPQ 1153
RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N:Alternate names: complement-3 receptor alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C:Accession: S00551; 159078
R:Pytela, R.
EMBO J. 7, 1371-1378, 1988
A:Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th
A:Reference number: S00551; MUID:88312584; PMID:3044779
A:Accession: S00551
A:Molecule type: DNA
A:Residues: 1-1153 <PYT>
A:Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983
A:Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.6%; Score 4447; DB 2; Length 1153;
Best Local Similarity 73.5%; Pred. No. 2.1e-302;
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARFGQSVVLOGSRVWVCAPOEIVAAANORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTEHPMTFQENAKGFQGVVQLGGTSVVAAPQEAQKAVNQTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVAVNMSLGLSLAATTSPQLLAGCTPVTQTCSENTYVKGFLCFLGSLNRQPOK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQQLLAGCTPVTQHCNKENTYVNGCYLFGSNLLRPPQ 136

QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHFRRMKELVSTIMEQKKSKTLFSLMQYSBEF 180
DB 137 FPEALRECQESDI VFLDGGSGINNIDFQKMKFVSTVMEQFKSKTLFSLMQYSDEF 196

QY 181 RIHFTKEQNPNPSLAKPITQLLQRTHTATGLRKVRRELPNITNGARKNAFKILFL 240
DB 197 RIHFTFNDFKRNPSRSHSPIQLNGRTKTASGIRKVRRELPHKTNAGRENAAKILVVI 256

QY 241 TDEKEKGDPLGYEDVPELDREGVIRVVLGFGDAFSEKSRQELNVTASKPRDHDVFOAN 300
DB 257 TDEKEKGDPLGYEDVPELDREGVIRVVLGFGDAFSEKSRQELNVTASKPRDHDVFOAN 316

QY 301 NFPEALTVQNLREKIFAIEGTQTSSSFEHEMSQEGFSAAITNSGPLLSTVGSVDWAG 360
DB 317 NFPEALTVQNLREKIFAIEGTQTSSSFEHEMSQEGFSAAITNSGPLLSTVGSVDWAG 376

QY 361 GVFLYTSKEKSTINTRVDSMDNDAYLVAAAILRNVRVQSLVGLGAPRYQHIGLVAMPR 420
DB 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLVAAAILRNVRVQSLVGLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWSEANVKGITGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 480
DB 437 ENFGTWEPTSHIKGSIQSGYFGASLCSVDMDADGNTNLILGAPHYVEOTRGQSVCP 496

QY 481 PRQARWOCDAVLGYEQGQPMWCFGAALTVDLVGNDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQCEALLHGDQHPWCFGAALTVDLVGNDKLTVAIGAPGEEDNRGAVYLF 555

QY 541 HGTSGSGISPHSHQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRQ 600
DB 556 YGASIASLSASHSHRIIGHFSPGLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRQ 615

QY 601 PVLRVKAIMEFNPREVARNVECDVWKGAGEVRVCLHVOKSTFDRLRREGIOQSV 660
DB 616 PVLRVKAIMEFNPREVARNVECDVWKGAGEVRVCLHVOKSTFDRLRREGIOQSV 675

QY 661 YDLALDSGRPHSAFVNETKNSRRTQVLGLTQTCETLKLQPNCEIPVSPVILRLNF 720
DB 676 YDLALDPVRSIRAFDEFKNTTRTQVFLMQKCEETLKLILPDCVDVSPVILRLNY 735

QY 721 SLVGTPLSAFGLNRPVLADAQRALTALFPFEKNCNDNICQDDLSITFMSLDCLVVG 780
DB 736 TLVGEPLRSGNLRPVLADAQRALTALFPFEKNCNDNICQDDLSITFMSLDCLVVG 795

RESULT 3

RWHLUC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150.95 alpha chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A:Residues: 1-1163 <COR>

A:Note: this revision to the sequence from reference A35543 includes the carboxyl end

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150.95 molecule

A:Reference number: A35543; MUID:90153306; PMID:2303426

A:Accession: A35543

A:Molecule type: DNA

A:Residues: 1-834 <CO2>

A:Note: this sequence has been revised in reference A36584

R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A:Reference number: S00864; MUID:88166645; PMID:3327687

A:Accession: S00864

A:Molecule type: mRNA

A:Residues: 1-755, 'L', 757-1163 <CO3>

A:Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:g487830

A:Note: part of this sequence was confirmed by protein sequencing

C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on

C:Genetics:

A:Gene: GDB:ITGAX; CD11C

A:Cross-references: GDB:119758; OMIM:151510

A:Map position: 16p11.2-16p11.2

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology

C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>

F:20-1107/Domain: extracellular #status predicted <EXT>

F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>

F:1108-1133/Domain: transmembrane #status predicted <TM>

F:1134-1163/Domain: intracellular #status predicted <INT>

F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Aen) (covalent) #status F

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151 <BRI>

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: sequence extracted from NCBI backbone (NCBIP:124326)

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.5%; Score 1085.5; DB 2; Length 1151;
Best Local Similarity 27.6%; Pred. No. 3.1e-67;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;

QY 1 FNLDTENAMTFOENARG-FOGSVQL---QGSRVVVGAPQEIIVAANQSGLSYQCDYSTGS 56
DB 1 FNVDKNSMTFGPVEDMFCYTVQYENEBEGKVLIGSLVGPQKRTGDVYKCPVGRGE 60

QY 57 CEP-IRLOVPEA-----VMSLGLSLAATTSPQQLACGPTVHTQCSNTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTVEKNMTFGSTL-VTPNPGGFLACGPLYAYRCGHLHYTTGI 119

QY 107 CFLFGSNLRQOPQKFPBALGCGQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMBOLKK 166
DB 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVIVLDGNSIYPWD-----SVTAFENLKKR 171

QY 167 -----SKTFLSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLGR-THTATGLRKVR 220
DB 172 MDIGPKQTQVGIVQYGNVTHEFNLNKYSSTEEVLVAAKKIVQGGRTMTALGTDATK 231

QY 221 ELFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRYVLGFGDAFR---- 276
DB 232 EAETEARGARGVKKNVVIITDGEH--DNHRLKKVIQDCDENIQRFSAIILGSGYRGNL 290

QY 277 -SEKSRDELNTVASKPRDHVFOANFEALKTQVONLRKIFALEGTQTGSSSFEHMS 335
DB 291 STEKFVEIKSASEPTEKHFNNVDELAVITVKTIGERIFALEATADGSAASFEMMS 350

QY 336 QEGFSAITSNGPLSLTVSGYDMAGGVFLYTSKE-----KSTF-INNTRVDSMDNAYLG 389
DB 351 QTGFSAHYSODWMLGAVGAYDNGTVMQKASQIIPRTTNNVESTKKNEL-ASYLG 409

QY 390 YAAAIILNRVQSL-VLGAPRYQHIGLVAMFRONTGMWESNANVKTQIGAYGASLCV 448
DB 410 YTVNSATASSGDVLYIAGOPRYNHTQVVIYRMEDGNKIKILQLSGEIGSYFGSILTTT 469

QY 449 DVDSNGSTDLVLGAPHY-----YEOTR-OGQVSVCLPRGQARQWQDA 492
DB 470 DIDKSDNTDILLVGAPMYMGTKEEEOGVVYVYALNQTFRFYQMSLEPIKOTCCSSRQHNS 529

QY 493 VLYGEOQOPWG-RFGAALTVLVDGVDGKLTDAVIGAPGEDNDRGAVLPHGTSGSGISPS 551
DB 530 CTTENKNEPCGARFGTAIAVKDLNLDGFNDIVIGAPLEDHGHGAVVIHG-SGKTIRKE 588

QY 552 HSQRIAGSKLSPLQYFGSLSGQDQITMDGLVDLTVGAQHVLRLRSQPLVRKAIMEF 611
DB 589 YAKRIPSGGDKTLKFFGQSIHGMDLNGDGLTDVTIGLGGAAALFWSRDVAVVYKVTMNF 648

QY 612 NPREVARNVECDNQVVKGEAG--EVRVCLHQV-KSTRRLREGQIQSVVTDALDSG 668
DB 649 EPNKNTQKKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVLDLSL 701

QY 669 RPHSRVAFNET-----KNSTRROTQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNLSL 722
DB 702 RQTSRSPFGTQERKVOBNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNL 753

QY 723 VGTPLSAGNLRPVLAAQRLFTALPFPFKNGCNDNICQDDLSITFSFMSLDCLVGGP 782
DB 754 T-DPENG-----PVLDDSLNSVHEYIPFAKDCGKCEKICSLSLHVAATTEKDLIIVRSQ 807

QY 783 RE-FNVTTVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 841
DB 808 NDKFNVSILTQKNSAYNRTTIVHYGNLVSFGIEAIOKD-----SCESN----- 853

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C;Accession: I45914

J;Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A;Title: Identification of putative ligand binding sites within the I-domain of integr

A;Reference number: A54402; PMID:94193647; PMID:7511592

A;Accession: I45914

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1170 <KAM>

A;Cross-references: GB:L25886; NID:9439695; PIDN:AAB59255.1; PID:9439696

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match

Best Local Similarity 27.5%; Pred. No. 3.3e-66;

Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

QY 1 FNLDTENAMTFO-ENARGFQSVQL---QGSRVVVGAPQEIIVAANQSGLSYQCDYST 54
DB 19 YNVGLPKAKIFSGPSSQEQFGYAVQQFINPKGNWLLVGSWSPGPKNRMGDVYKCPVDLST 78

QY 55 GSCEPIRLO-----VPVEAVNMSLGLSLAATTSPPQLACGPTVHTQCSNTYVKG 107
DB 79 TTCEKLNLOTSTMSMNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAQCGSQYTTGVC 138

QY 108 FLFGSNLRQOPQKFPBALGCGQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLK-- 165
DB 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVCDENSIYPWD--AVNFKLFVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLL-----GRTHATGLRKVR 221
DB 195 PTKQTMGLIQYANNPRVFNLTNFTKSD---EMIKATSQTFQYGGDLTNTFKAIQVARDT 251

QY 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRY---VLGF-GDAFR 276
DB 252 AYSTAAGGPGATKVMVVTVDGESH--DGSKLKAVIDQCNDKNTLRFGLVGLYLNRLD 310

QY 277 SEKSRDELNTVASKPRDHVFOANFEALKTQVONLRKIFALEGTQTGSSSFEHMSQ 336
DB 311 TKNLKIEIKAIASIPTEHFHFNVSDEADLLEKAGTIGEQIFSGIEGTQVQ-GDNFQHEMSQ 369

QY 337 EGFSAAIT--SNGPLLSTVGSVDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGVA 391

Db 370 VGFAEYSPONNMLGAVGAYDWSGTVVQKTPHGLIFSKQAPEQLQDRNHSYLGYS 429
Qy 392 AAILLRNVQSVLIGAPRYQHIGLVAMFRONTGWNESNANV-----KGTQIGAYFCASL 445
Db 430 VASISTGNSVHFVAGAPANTYTGQIVLYSVN-----ENGNTVVIQSQRGDDIGSYFQSVL 484
Qy 446 CSVDVDSNGSDLVILGAPHYVEQTR--GGQVSVCPPLPRGQARWQCDVAVLYGQGGPWG 503
Db 485 CAVDVNDKTTIDVLLVGAPMYMNDLKKEGRVYLFTTKG-ILNWH--QFLEGPNLENA 541
Qy 504 RFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSORIGAS--KL 561
Db 542 RFGSAIALSDINMDGDFNDVVGSPLENQNSGAVIYNGHEGM-IRLRYSKILGSDRAF 600
Qy 562 SPRLOYFQCSLGGDLTMDGLVDLTGCAQGHVLLRSQPLVLRKAIMFNPVARNVP 621
Db 601 SSHLOYFQSRSDGVGDLNGDSITDVSAGFAQVQVQLWSQSADVSQASFPFKKI--TLL 658
Qy 622 ECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVYDIALD----SGRPHSRVAFN 677
Db 659 NKNAEI-----KLKLCF----SAKPRPTNQNNQVAIVYNTIDEDQFSSRSVLSGLFK 707
Qy 678 ETQNSTRTQVGLGTQCE--TLKLQPLNCIEDPVSIVLRNPSL--VGTPLSAGNL 733
Db 708 ENNERCLOKTMVSAQRCSEYIIHIOEPS---DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVLAEDQRLFTALFPFCKNCGNDNICDDLSITF-----SFMSLDCLVVGPPREFNVTV 789
Db 757 NPALEYSTVKTVPISFPHKQDGGDVCISDLVNLVQQLPATQQOPFIVSNQNLRTFSV 816
Qy 790 TVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKSTS 848
Db 817 QLKNNKESAYNTEIIVDFSENLF-----ASMSMPVDGTEVTCQIASSQKSVT 864
Qy 849 CSINHIPIPEENSEVTNFTFDVDSKASLGNKLLKANVTSENMPRTNKTETQLELPVKY 908
Db 865 CNVGYPALSKQKQVTFTFNDFNLQ-NLQNAQASISFRALSESQEBENMADNVLNLSLLY 923
Qy 909 AVYVMTVSHGVSTKYLNTASENTSRVMQHOVQVSNLQOR-----SLPTSLVFLV 958
Db 924 DAETHIT-RSTNINFEVSLDGNSSVV-HSFE--DIGPKFISIKVTGVSVPVMSA--- 976
Qy 959 PVRLNQTVIWRDPQVTFSEN---LSSTCHTKE-----RLPSSHDFLAE- 998
Db 977 -----SVLIHPQYTKDKNPLMYLTGVHTDQAGDISCEAEINPLKIQTSSTSSVSFKSEN 1030
Qy 999 LRKAPVNCIAVCORIQCDDIPFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABI- 1057
Db 1031 FRHIKELNCRATASCNIMCWLRLQVKGVEYFLNVSTRIWNGTFAASTQTQVLTAAABID 1090
Qy 1058 LFENDSVFTL-----LRGQAFVRSQETKVEPPE-VPNPLPIVSGSVGGLLLALITA 1110
Db 1091 TYNPOIYVIENTVTIP-----LTIMPKHEKVEVPTGVIVGSVIAGILLALLALVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFFKRYEKM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:G473098; PIDN:CAA82877.1; PID:G473099
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
P:169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;

Best Local Similarity 27.8%; Pred. No. 3.2e-85;

Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIIVAAHQRSGLYOC--DYST 54
Db 27 YNVGLPKAKIFSGPSSEQFGYSVQQLTNPQGNWLLVGSWSPGPEPNRMDGVYKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQLLACQPTVHQTCSNTYVKGLC 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCPLWAHQCGNOYATGIC 146
Qy 108 FLFGSNLRQPOQ---KPEALRCQPEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQL 164
Db 147 ----SDVSPDFQFLTSPAPVQACPSL-VDVVVVCDESNIYP--WEAVKNFLVKFVTGL 199
Qy 165 K--KSKTLFSLMOYSEBFRIHFTKBFQNNPNRSLIKPITQLLG-RTHTATGLRKVRE 221
Db 200 DIGPKKTOVALIOYANEPRIIIFNLNDFETKEDMVQATSETROHGGDLTNTFFRAIBFARDY 259
Qy 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVIPLELDREGVIRY---VLGF--GDAFR 276
Db 260 AYSQTSGRPRGATKVMVVVTDGESH-DGSKLKTVIQCNDDIELRFGIAVLGYLNRNALD 318
Qy 277 SEKSRQELNVTASKPRDRHVFQANNFEALKTQVQNLREKIFAIEGTCTGSSSEFHEMSQ 336
Db 319 TKNLKIKETKAIASITPTERYFFNVADAEALLEKAGTIGEQIFSIETGVQ--GDNFQWEMAQ 377
Qy 337 EGFSA--AITSNGPLSLTVSGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
Db 378 VGFSAADYAPQNDILMLGAVGAPFMSGTLVQETSHKPVIFPKQAFDQVLDQRNHSFLGYS 437
Qy 392 AAILNRVOSLVGLAPRYQHIGLVAMFRONTGWNESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHVFAGAPRANVTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYVEQTR--GGQVSVCPPLPRGQARWQCDVAVLYGQGGPWG 505
Db 495 VDVKDQITDVLVVGATPYMNDLKKEGKLYLFTITKILNQHO---FLEGPEGTGNARF 551
Qy 506 GAALTVLGVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSORIGASLSPR- 564
Db 552 GSAIAALSDINMDGDFNDVVGSPLENQNSGAVIYNGHEGM-IRTKYSOKILSNGAFRR 610
Qy 565 -LQYFGOSLGGQDLTMDGLVDLTGCAQGHVLLRSQPLVLRKAIMFNPVARNVPFEC 623
Db 611 HLQFFGRSLDGYGLNGDSITDVSIGALQVQLWSQSIADVAIEALFTP----- 660
Qy 624 NDQVVKKEAGEVRVCLHVQKSTRDLREGQIQSVVYDIALD----SGRPHSRVAFNET 679
Db 661 -DKITLLNKQAKITLKLCFRAEFAPAGQNNQV--AILFNMTLDADGHSSRVSRTSGVFPREN 717
Qy 680 KNSTRROTQVGLGTQTCET--LKLQPLNCIEDPVSIVLRNLSFLVGTPLSAFGLNLRPVL 737
Db 718 SERFLQKNVNVQVKCESEHHSIQKPS---DVVNPDLRLVDLSLENPQTS-----PAL 768
Qy 738 AEDAQRLLTALFPPEKNCNDNICQDDLSI-----TFSPMSLDCVLVGGPREFNVTV 788
Db 769 EAYSETVKVFSIPFYKEGSDGICISDLILDVQQLPAIQTSF-----IVSNQNKRLTFS 823
Qy 789 VTVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKST 847
Db 824 VILKNRGESANTYVLAEPFSENLF-----ASFSPMPVDGTEVTCVGSQKSV 871
Qy 848 SCNHNPIFPENSEVTNIFTVDVSKASLGNKLLKANVTSENMPRTNKTET--FQLELP 905
Db 872 TCDVGYPALSKQKQVTTTINDFNLQ-NLQNAQAINFQAFSESO--ETNKADNSVSLTIP 928
Qy 906 VKYAVYVMTVSHGVSTKYLNTASENTSRVMQHOVQVSNLQOR-----SLPISLV 955

```

Db 929 LLYDAEHLT-RSTNINFIYSSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVMA 984
Qy 956 FLV-----PVRINQTVIMDRPQVTF-SENLS 980
Db 985 LVTHIQYTKENPLLYLTGIOTDQAGDISCTAEINPLKLPHTA-----PSVSFKNEFR 1040
Qy 981 STCHTKERLSHSDFLAELRKAPVNVGSIACVORIQCIDIFFGQIBEFNATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMAKAEYFINVITRVNRT 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDVSFTLLPQGGAFVRSQTETKVEFVNPPLPLIVGSSVG 1100
Db 1081 FPASTFQTVQTAANEIDTHNPQFVIEENAVTIPLMIMKPTKAEVPT---GVIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFKQYKDM 1125
Db 1139 GILLLLAMTAGLWLGFKQYKDM 1163

RESULT 10:
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 21-Jul-2000
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <PAK>
A:Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R:Cattell, B.; Parentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIc*, GPIIa an
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16,'V',18-21 <ZUT>
A:Cross-references: GB:L24121; NID:G400342; PIDN:AAA16619.2; PID:G4583535
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.8%; Pred. No. 5.2e-65;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARCEQSVVQL---QGSVVVVCAPQEIIVAAQNRGLYQC--DYST 54
Db 30 YNVLPPEAKIFSGPSSEQFGYAVQQFINPKGNLLVGSPPSWGPFENRMGDVYKCPVDLST 89

```

```

Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPOLLACQPTVHQTCSENTYYKGLC 107
Db 90 ATCEKLNLOSTSIPIVNTKMTNMSLGLILTRNMGSGFLTCPLQAQCCNQYTTTGV 149
Qy 108 FLFGSNLRQOPQKFPALRGCPQEDSDIAFLIDGSGIIPHDPFRMKELVSTIMEQK-- 165
Db 150 SDISPDF-QLSASFSPATPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKPVQGLDIG 205
Qy 166 KSKTFLSLMQYSEEFRIHFTKEFQNPNSRSLIKPITQLLG-RHTATATGLRKRVRLEFN 224
Db 206 PKTQVGLIOYANNPRVFNLTNTYKTEEMIVATSTQYGGDLTNTFCAIOIYARKYVS 265
Qy 225 ITNGARKNAFKILFLTGDKEFGDPUGYEDVPELDEGVIRY---VLGF--GDAFRSEK 279
Db 266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNDHNLIRFGIATVGLYLRNALDTKN 324
Qy 280 SRQELNTVASKPRDHVFOANNFEALKTQVONLREKIFAIEGTQTQTCSSSEFHEMSQEGF 339
Db 325 LIKEIKAIASIPTEYFFNVSDAEALKEKAGTLGEQIFSEGTVOG-GDNFQEMSQVGF 383
Qy 340 SAAITSGNP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGAAAI 394
Db 384 SADYSQNDILMLGAVAGMGSTIVQKTSHGHLIFPKQAFDQILQDRHSSVYLGTSVAA 443
Qy 395 ILNRNVQSLVGLAPRYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFVAGAPRANTYTGQIVLYSVN-----ENGNITVIAHRGDDGIGSYFGSVLCSV 498
Qy 449 DVDSNGSTDLVLIGAPHYEOTR--GGQYSVCLPRQGRARWQCDVAVLGEQOQPWGRFG 506
Db 499 DVDKDTITDVLVVGAPMYSDLKKEGRVYLFITKKGILGQHQ---FLEGPEGIENTRFG 555
Qy 507 AALTVLGVNGDKLTDVAIGAPEEDNRGAVLVFCHGTSGSGISPSHSORAGS--KLSR 564
Db 556 SAIAALSDINMGDFNDVIVGSPLENQNSGAVIYNHGHGT-IRTKYSQKILGSDGAFRSH 614
Qy 565 LQVFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQPLVLRVKAIMEFNPREVARNYFECN 624
Db 615 LQVFGRLDGYGDLNGDSITDSIGAFQVQVQLWSQSIADVAIEASTPEKI--TLVNKN 672
Qy 625 DQVVVKEAGEVRVCLHVOKSTRDLRREGIQSVVITYDLALD-----SGRPHSAVRVNETK 680
Db 673 AQII-----LKLCP---SAKFRPTKQNNQVAVIYNTITLADGFSRVRTSRGLPKENN 721
Qy 681 NSRRTOTVGLTLTOTC--ETLKLQPNCTEDPVSPIVLRNLNFSLVGTPLSAFGLNLPVLA 738
Db 722 ERCLQKNMVVQAQSCPEHIYIQEBS---DVVNSLDLRVDSLENPGTS-----PALE 772
Qy 739 EDAQLRFTALFPPEKNCNDNICQDDLSITF---SFMSLDCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEDGLCISDLVLDVRQIPAAQOPPFIVSNQNKRLTFSVTLKNK 832
Qy 795 GEDSYRTQVTFPPPLDLSVRKUSTLQNSQRSWRLACESASST-EVSGALKTSTCSINH 853
Db 833 RESAYNTGIVDFSENLF-----ASFSLPVDGTEVTCQVAAASQKSVACDVGY 880
Qy 854 PIFPENSEVTENTFDVDSKASIGNKLLKANVTSENNMPRTNKTFFOLELPVKYAVYMW 913
Db 881 PALKREQQVTFITFDFNLQ-NLQNASLUSFOALSEBENKADNLVNLKIPILYDAEI- 938
Qy 914 VTSHGVTKYLVNFASENTSRVMQHOYVSNLQOR-----SLPTSLVFLV----- 958
Db 939 ---HLTRSTININFEYISSDGNVPSIVHSFEDVGPKEIFSLKVTGTSVPVSMATVIHIPQ 995
Qy 959 -----PVRINQTVIMDRPQVTF-SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGQT---SSSVSPKSENF--HTKE- 1047
Qy 989 LSHSDFLAELRKAPVNVGSIACVORIQCIDIFFGQIBEFNATLKNLSFDWIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCLWKDVMHMKGBYFVNVTTRINWGTTFASFTQT 1091

```

Qy : 1049 LLIVSTAEI-LFNDSTVPTLLPGGAFVRSOTETKVPFVFPNP-----LP--LIVGSSVG 1100
Db : 1092 VQLTAAAEINTYNEPIYVI-----EDNTVTIPLMIMKPKDEKAEVPTGVIIGSIIA 1141
Qy 1101 GLLLALITAAALYKLGFFKROYKDM 1125
Db 1142 GILLALVALVAILKLGFFKRYEKM 1166

RESULT 11
A35854
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1990 #sequence.revision 13-Sep-1991 #text_change 20-Sep-1999
C:Accession: A35854; S11243
R:Ignatius, M.J.; Lorge, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
A:Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A
A:Reference number: A35854; MUID:90338125; PMID:2380249
A:Accession: A35854
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1180 <GN>
A:Cross-references: GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
F:170-345/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 17.8%; Score 1049; DB 2; Length 1180;
Best Local Similarity 27.3%; Pred. No. 1.2e-64;
Matches 344; Conservative 1200; Mismatches 476; Indels 240; Gaps 48;

Qy 1 FNLDTENAMTFOENARG-FGQSVVQL---QGSRRVVVGAPOEIVAAANQRGSLGYOCDYSTGS 56
Db 29 FNDVKNMSFGSGVEDMFGYTVQYENESGKWVIGSLVGPQPKARTGDVYKCPVGRER 88
Qy 57 CEP-IRLQVPVEA-----VNMSLGLSLAATTPSPQLLACGPTVHQTCSNTYVYKGL 106
Db 89 AMPCVKLDLPNTSIPNVTBIKENMTFGSTL-VTNPNGGFLACGPLYAVRCGHLHYTTGI 147
Qy 107 CFLFGSNLRQOPQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMQLKK 166
Db 148 CSDVSPPTQVNVNAP--VQECSTQ-LDVIIVLDGNSNIYP-----WESVIAFNLDLKR 199
Qy 167 -----SKTFLSLMOYSEEFRIHFFKFFQNNPRSLIKPITOLLG-RTHATATGLRKVVR 220
Db 200 MDIGPKOTQGVIVQYGENVTHFENLNKYSSTEEVLVAANKIGRGGGLQTTWALTGIDTARK 259
Qy 221 ELFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDREGVIRY---VLFGGDA--F 275
Db 260 EAFTEARGARRGVKVMVIVTDGESH-DNYRLKQVIQDCEDENIQRFPSIALGHYNGNL 318
Qy 276 RSEKSRQELNTVASKPRDRHVQFQANFEALKTQVQNLREKIFAIEGTQTGSSSFHEMS 335
Db 319 STEKFVEIKSIASEPTKEKHFNVSDSLALTVIKALGERIFALEATADQSAASFEMS 378
Qy 336 QEGFSAAITNSGPNLSTVGSVDWAGGVFLVTSKEKSTFINMT--RVDSMDND---AYLGY 390
Db 379 QTGFSAHYSQDWMGLGAVGDMNGTVMQKANQWVPHNTTQTETPAKNEPLASLYGY 438
Qy 391 AAAILRNVRQSLVLGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGALSCSDV 450
Db 439 TVNSATIPGDVLYIAGQPRYNHTQGVVIKYMEDGNINILQTLGGEQIGSYFGSVLTIDI 498
Qy 451 DSNGSTDLVLIGAPHY-----YEOTR-GGVSVCPPLPRGQARWQCDAVL 494
Db 499 DKOSYTDLLVGAPMYMGTEKEEQGVVYAVNQTREYQWMSLEPIRQTCCSLKDNCSCT 558
Qy 495 YGEOGPWG-RFGAALT/VLGDVNGDKLTVDAIGAPGEDNKGAVLYFHEGTSGGISPSHS 553
Db 559 KENKNEPCGARFGTAAIAVDNLNDVGNFVVIAPLEDHDAGAVIYTHG-SGKTIREAYA 617
Qy 554 QRIAGSKLSRLQYFGQSLSGQDQLTMDGLVDLTVGAQGHVLLRLSRQVPLRVKAIMFNP 613

```

Db      618  QRIPSGGDKTKLFFQSGTHGEMDLNGDGLTVITIGLGGAAALFWARDVAIVKVTWNFEP 677
Qy      614  REARVNVFECNDQVWKGKEAG--EVRVCLHVQ-KSTRDLRLEQGIQSVWYDLDALSGRP 670
Db      678  NKVNIQKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ---YRVLDSLRLQ 730
Qy      671  HSRVAVNET-----KNSTRQTVGLGLTOTCETLKLQLPNCI-----BDPVSPV 715
Db      731  ISRSFFSGTQERKIQRNITVRESE-----CIRHSFYMLDKRHFQDSVR 773
Qy      716  LRLNFSIVGTPLSAFGNLRVLAEDAORLFTALFPEKQCGNDNI CODDLSITFSFMSLD 775
Db      774  VTLDNFILTPENG-----PVLDDALPNSVHEHIFPAKOCGNKERCISDUTLNVSTTEKS 827
Qy      776  CLVVGGPREFNVTVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACES 834
Db      828  LLIVKSQHDKNVSLTVKNGDSAYNTRTVVQHSPLNIFSGIEIQKD-----SCES 879
Qy      835  ASSTEVSGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLL-LKANVTSENMP 893
Db      880  N-----QNITCRVGPFLRAGETVTFKIIIFQFNSTHLSLSENAIHLSATSDSEEPLE 930
Qy      894  RTNKTPEOLELPVKYAV---YMWVTSHGYST-----KYLNFTASENTRVWQHOYQV 942
Db      931  SLNDNEVNISIPKYEVGLQFYSSASEHHSVAANETIPEFINST-EDIGNEINVEYTI 988
Qy      943  SNLGQSRSLP--LISLVF-----LVPVRLNQTVMD-----RP----- 971
Db      989  RKGHGHPMPELQISLFPNLTADGYPVLYDIG-----WSSDNVNCRPSLEDPPGINS 1042
Qy      972  --QVTF-----ENLSSTCHTKERLPSSHDFLAELRKAPVNVNCISIAVCORIQCDI 1019
Db      1043  GKXWITISKSEVLKRGTIQDCSSTC-----GVATITCSLLPSDLSQVNV 1085
Qy      1020  -----PFFGIQIEF---NATLKNLSFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQ 1070
Db      1086  SLLLWKPTF-IRAHFSLNLTLRGELK-----SENSLITLSSN----- 1123
Qy      1071  GAFVRSOTENKVPFEPNPLPL--IVGSSVGGLLILALITAALYKLGPFKRYQKDMWSE 1128
Db      1124  ---RKRELAIQISKDGLPGRVPLWVILLSAFAGLLMLLILALWKIGFPKRPPLKKQWEK 1180

RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N:Alternate names: Integrin alpha-4
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
C:Accession: A41131; S16742
R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.
J. Cell Biol. 115, 1149-1158, 1991
A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine ly
A:Reference number: A41131; MUID:92064645; PMID:1840602
A:Accession: A41131
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1039 <NU>
A:Cross-references: EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g51485
C:Superfamily: Integrin alpha-4 chain
C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 666; DB 2; Length 1039;
Best Local Similarity 22.8%; Pred. No. 5.9e-38;
Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48

Qy      1  FNLDENAMTFQ-ENARGFQSVV-QLQGRS--VVVGAPORIVAAN----ORGSLYQCDY 52
Db      41  YNLDPENALLYQPSGTLFGYSVVLHSHGSKRWLIIVGAPTASWLSNVAASVNVPGAIYRCGI 100
Qy      53  STG---SCEPIRLQVP-----VEAVNMSLGLSLAATTSP-POLLACG---PTVHQ 95

```


QY 343 ITSNGPLSTVSGYDWAGGVFLY---TSKEKSTFTINMTTRVDSMDNDAYLGYA--AAIILR 397
DB 212 YTKDLIVNAGPSSWTSLSFYNTITNKYKAFDKQKQVXF--GSYLGSVCGAGHPRS 268
QY 398 NRVSGLVGAPRYQHIGLVAMPQNTQMGWESNAV---KGTQIGAYFGASLCSVDVDSN 453
DB 269 QHTTEVVGAPQHEQIGKAYIF---SIDXELNIIHEMKKGLKSGYFGASVCAVDLAD 324
QY 454 GSTDLVLIGAPHYEQTRGGGVSCPLPRGQARWQC--DAVYGEQGPWGRFGAALTVL 512
DB 335 GFSDL-LVGAPMQSTIREGRVFY-INSGSAVNMAMETNLVGSCKYA-ARFGESIVNL 381
QY 513 GDVNGDKLTVAIGAPEGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSL 572
DB 382 GDIDNDGFEDVAIGAPEGEDDLQGAIIYVNGRA-DGISSTFSQRIEGLQISKLSLMPGQSI 440
QY 573 SGGQDLTMDGLVDLTGVA--QGHVLLRSQPVLRVKAMENPREVARNVFECNDQVKG 630
DB 441 SGQIDADNNGYVDVAGAFRSDSAVLLRTRPVVIDASLS-HPESVNRKTFDC----- 492
QY 631 KEAGEVRVCLHVOKSTRDLREGQIQSVVTVYDLALDSGR-----PHSRAVENETKNSTR 685
DB 493 VENGWPSVICIDLTCSYKGEKVPYIVLFYNMSLDVNRKAESPFRFYFSSNGTSDVITG 552
QY 686 QTVLGLTQTCETLKLQPNCTIEDPVSPVLRNLSLVGTPLS-----AFGNLRPLAED 740
DB 553 SIQVSSREANCRTHQAFMRKDVDRDILTPIQIEAAVHLGPHVIVSKRSTEEFPLQILOQK 612
QY 741 AQR-LFTALFPKKNCGNDNICQDLSIT--PSPM-----SLDCLVVGQPREFNVTVYRN 793
DB 613 KEKIDMKKTINFARFCAHEN-CSADLQVSAKITGFLKPHENKTYLAVGSKMTLMNLNLFN 671
QY 794 DGEDSVRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASSTEVSGALKSTCSINH 853
DB 672 AGDAYETTLHVKLPGVYFIKILEEK-----QINCE---VTDNSGVVQ-LDCSIGY 721
QY 854 PIFPENSEVTFNITFDVDSKASLGNKLLKANVTSEN--NMPTNKTBFQLELPVKYAVY 911
DB 722 IYVDHLRSRDISFLDLVSSRAEEDLSITVHATCENEEMDNKHSRVTVAIPUKYEVK 781
QY 912 MVVTHSGVSTKYNLTAFASNTSERV-----MOHOYVSNLQORSLP-ISLVLVPLVRLNQT 965
DB 782 LTVHGVFNPTSVFYNSDNEPETCMVEKNNLTFFHINTGSMAPNVSVSEIMVNSFS-- 839
QY 966 VIWDAPQVTFSENL-----SSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019
DB 840 -----PQTDKLFNLDVQTTTGECHFNQ-----RVCALEQ--- 871
QY 1020 PFFGIQEEFNATLKGNLSPDWYIKTSHNHLIVSTAE-----ILFN-----DSVFT 1065
DB 872 -----QKSAMQTLKGVRF--LSKTDKELLYCIKADPHCLNLFNFGMGESGKEASVHI 923
QY 1066 LLPQGAFAVRSQTEKVEPFEV-----PNP----- 1090
DB 924 QLEGRPS-ILEMDETSALKFEIRATGFPPEPNRVIELNKNDENVAHVHLEGLHQRPKRYF 982
QY 1091 -LPLIVSGSVGLLIALITALYKLGFFKROVKDMMSE 1128
DB 983 TIVIISSLLGLLIVLLISYVNMKAGFFRQVKSILQE 1021

RESULT 14

I58409

integrin alpha-9 chain precursor - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000

C:Accession: I58409; A49459

R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah

Oncogene 9, 611-619, 1994

A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small

A:Reference number: I58409; MUID:94119603; PMID:8290272

A:Accession: I58409

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1035 <RES>

A:Cross-references: GB:D25303; NID:G464180; PIDN:BAA04984.1; PID:G533327

R:Palmer, B.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.

J. Cell Biol. 123, 1289-1297, 1993

A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel par

A:Reference number: A49459; MUID:94064789; PMID:8245132

A:Accession: A49459

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 30-1035 <PAL>

A:Cross-references: GB:L24158

C:Superfamily: integrin alpha-4 chain

C:Keywords: Glycoprotein; metal binding; transmembrane protein

F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7% Score 630; DB 2; Length 1035;

Best Local Similarity 26.5%; Pred No. 1.9e-35;

Matches 23; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

QY 331 EHEMSQEGFSAATISNGPLSTVSGYDWAGGVFLYTSKEKSTFTINMT-RVDSMDNDAYLG 389

DB 190 EHGSCQAGIAGFFTEELVVMGAPGSGFYWAGTIKVLNLT-D-NLYLKLNDDEVIMNRRYTYLG 248

QY 390 YA--AAIILNRVQSLVLGAPRYQHIGLVAMFR--QNTGMWESNAVVKGTQIGAYFGASL 445

DB 249 YAVTAGHFSPSTTIDVVGGAPODKGKGVYIFRADRRSGTLIKIFQASGKMGSGYFGSSL 308

QY 446 CSVDVDNSGSTDVLIGAPHYEQTRGGGVSCPLPRGQARWQCDVLYGEOQOPWGRF 505

DB 309 CAVDLNGGSLD-LVGAPMFSEIRDEGQVTVY-INGNGALEE-QLALTGD-GAYNAHF 364

QY 506 GAALTVLGDVNGDKLTVAGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRL 565

DB 365 GESTASLDDLNDGFPDVAIGAPKEDDPAGAVYIYHGDAG-GIVPQYSMKLSGGKINPVL 423

QY 566 QYFGQSLSGGQDLTMDGLVDLTGCA--QGHVLLRSQPVLRVKAMENPREVARNVFEC 623

DB 424 RMFGQISGGIDMDGNGYDPVTVGAPMSDSYVLLRARPVITVD-VSIFLPGSINITAPQC 482

QY 624 NDQVVKGKEAGEVYV-CLHVQ-----KSTRDLRREGQIQSVVTVYDLALDSGRPHSRV 675

DB 483 HD-----CQPVNCLNVTTCFSFHGKHVPEI---GLNYVLMAVAKKEKGMPRVY 531

QY 676 F---NETKNSTRRTQVLGLTQTCETLKLQPNCTIEDPVSPVLRNLSF-----VGTPLS 728

DB 532 FVLLGETMGQVTEKLQTYMEETCRHYVAHVKKRVQDVISPIVPEAAVYSLSEHVTEBER 591

QY 729 AFGNLRPVL-----AEDAQRLEFALFPPEKNGNDNICQDLSITFSFM--SLD-----CL 777

DB 592 ELPLPLTPVLRWKKGKIAQKNQTV---PERNCSED-CAADLQLOGKLLSSMDKLTLYL 647

QY 778 VVGGPREFNVTVYRNDCGDSYRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASS 837

DB 648 ALGAVKNISLNLISNLGDDAYDANVSFNVSRLEFF-----INMQKEEMGISCELLES 701

QY 838 TEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNK 897

DB 702 DPL-----KCSVGFPPMRSKSKYEFVIFDTSHLS--GEEVLVSFVITAQSG--NTER 750

QY 898 TE-----FQLELPVKYAVYVWVTS-----HCVSTKYNLFTASENTS-----RVNQHQY 940

DB 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFVYGESVDAANFIQLDLECHFOPIINTL 810

QY 941 QVSNLQORSPLTSLVFL-VPVRLN-----QTVIMDRPQVTFSENLSSTCHTKER 988

DB 811 QYNTGPTSLPGSSVSISFPRNLSSGGAEHFVQEMVVGQKNGCSFOKNPTPCIIPEQ 870

QY 989 LPSHSDFLAELRKA-----PVVNCISIAVCQRIQCDIPFFGIQEEFNATLKGNLSPD 1039

DB 871 ENIFHTIFAPFTKSRKVKLDCEKPGISCLTAHCN-----FSALAKESRTI-----D 917

QY 1040 WYIKTSHNHLIVSTAEILFNDSVFTLLPGQAGFVRSQTEKVEP----- 1084

Db 918 IV-----MLLNT-EILKXSSSVIQ-----FMSRAKVVDPAALRVVEIAHGNPBEV 962
QY 1085 ---FEVFN---PLPLTVG-----SSVGGLLLLALITAALYKLGFFKROYKQKDM 1126
Db 963 TVVFEALHLEPRGYVGMIIAISLLVGLIFLLLAALLWKMGFFRRRYKEII 1015

RESULT 15
T31437
Integrin/alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.5% Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 2.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAAITSNGB--LLSTVGSYDAGGVFLYTSKEKSTFINNTRVDS---DMNDAYLG 389
Db 181 QAGFSGIIFDSNALVWAGPGSYLQGIYVQSLNRSV-VQATGESNTGTSFDSNSYRG 239
QY 390 YAAAI--ILNRVQSILVIGAPYQHI-GLVAMFRQNTGMESNANVKGITQIGAYFGASLC 446
Db 240 YSLALGDFNGGVQDYVVGTPRAESLMGLVAIFDQNLQFN--QVMGTQIVAYFGYSVT 296
QY 447 SVDVDNSGTDLVLCAGPHYEQTRCGQVSVCPFRQQRARWQCDVLYGEQ----- 498
Db 297 VVDI--NNDYDILLVGAPEYMDGPAIQ-----RWEAGAVYYLQNPDPVGPQA 343
QY 499 -----GQPMGRFGAALTVLGVNGDKLTDVAIGAPCEDNRCGAVYLFHGTSGS 546
Db 344 SNRLSSTLIGGQIRSGELSGIASIGDSNQDGFNDVAIGAPYEGDDAGAVTYHG-SAN 402
QY 547 GISPSHSQRIAGSKLS-PRLOYFGQSLSGQDLTMDGLVDLTVGAQ--GHVLLRSQPYL 603
Db 403 GLKSTPAQVLTPTLGHSGITTFGSLQGGQMDKNKYPDLLVGAESANTAVLIRTPV 462
QY 604 RVKAMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRLRGQIQSV----- 658
Db 463 SLDATLNTPE-----IGINLENTYE--LADGTWVTSFIAMT 497
QY 659 -----VTYDLALDSG-RPHSRVAFNETKSTRQTQVLGL-TQTCETLKQ 702
Db 498 CFTYTGNYLPDHDIDSITVTVDSGIIANRRAMPVNDMSEITKTRRLAVSTQFCPLRAY 557
QY 703 LPNCIEDPVSPVLRNLNPLSVGTPLSAFGN-----LRPVLAEADAQRLFTALFPPEK 753
Db 558 VGNISIEDKLTPIKVTLQYDL-----NNDSERLQPHLEIPIDMATMTOTKQVSIQN 609
QY 754 NCGNDNICQDDISITFSFMSLCLVVGGRPFNVTVVRNNGEDSYRTQVTFPPFLDLSY 813
Db 610 NCVN-NICIPDLQVTVT-PNLPNIVIGTQELTLDVSLNRRGEDAFQSSLSVYPLGLOF 667
QY 814 RKVSTLQNSQSRWELACESASSTEVSGALKSTCSINHPIFPEN-----SEVTENIT 867
Db 668 VRL-----ERKANMDFSVTGSQSD-----LRIITCDTGNPMVGKNIILEFGLTLSTFQVS 717
QY 868 FDVDS-----KASLGNKLLKANVTSENPNRNKTEFQELPVPKYAVYVMVTVSHGVSTK 922

Db 718 GDKDSIEFYFKAESSENS--EDPNTLENNELNMTVPVTVDCTKLLLSASYPEIVMYSTQED 775
QY 923 YL-----NFTASENTSRVMOHOYOV-----SNLQORSPLIS-----LVFLVPV 960
Db 776 YVPPPPPAKNASEADIGMEVM-HLYEVNRTSSNAGEVSLNIQWPKQNEDEGEYLFYLLGI 834
QY 961 RLNQTVIMDRPQ-----VTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVV 1005
Db 835 MTEEGVTCQLTQGRKANPEGVKLEPSTKAKLSNSTTVQSVGRKRREPEVAEALAQTDN--VI 892
QY 1006 NCSTAVCORIQCDIPFFGIQEFNAT-----LKNLSFDVYIKTSHNHLIVSTAEILF 1059
Db 893 YCADSCVLINCTI-----DEINASKSVVRILGRF--W-----ERTF 928
QY 1060 NDSVFTLLPGOGAFVRSQTEKVE--PFEVNP-----LP----- 1092
Db 929 QKAVSELTVPVQATIASSASAANKTIPYINPLPRDFSSTKASTLVTTTEELVPPVPTIAM 988
QY 1093 -LIVGSSVGGLLLLALITAALYKLGFFKX-----QYKDMMS--EGGPP 1132
Db 989 WIIIVSVLGGIILLIILGLWKCGFFERKKPGEEKEYPVASADKDGPP 1038

Search completed: November 25, 2003, 14:21:55
Job time : 19.9579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNDLTENAMTFQENARGFGQ.....FKRQYKDMMSGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5823.5	99.1	1152	1	ITAM_HUMAN
2	4447	75.6	1153	1	ITAM_MOUSE
3	3446	58.6	1163	1	ITAX_MOUSE
4	3388	57.6	1162	1	ITAD_HUMAN
5	1536.5	26.1	1170	1	ITAD_HUMAN
6	1516.5	25.8	1163	1	ITAL_MOUSE
7	1136.5	19.3	1167	1	ITAE_MOUSE
8	1128	19.2	1179	1	ITAE_HUMAN
9	1085.5	18.5	1151	1	ITAI_HUMAN
10	1079.5	18.4	1189	1	ITAH_HUMAN
11	1071	18.2	1170	1	ITAE_BOVIN
12	1057	18.0	1178	1	ITAE_MOUSE
13	1054	17.9	1181	1	ITAE_HUMAN
14	1049	17.8	1180	1	ITAE_MOUSE
15	1042.5	17.7	1167	1	ITAG_HUMAN
16	666	11.3	1039	1	ITAA_MOUSE
17	642	10.9	1038	1	ITAA_HUMAN
18	630	10.7	1035	1	ITAA_MOUSE
19	594.5	10.1	1032	1	ITAA_XENLA
20	571.5	9.7	1066	1	ITAC_CRISP
21	567.5	9.7	1053	1	ITAC_MOUSE
22	555.5	9.4	1053	1	ITAS_MOUSE
23	551.5	9.4	1034	1	ITAV_CHICK
24	542.5	9.2	1050	1	ITAA_XENLA
25	539	9.2	1044	1	ITAV_MOUSE
26	534.5	9.1	1130	1	ITAE_HUMAN
27	533.5	9.1	1072	1	ITAE_CHICK
28	532	9.0	1049	1	ITAS_HUMAN
29	531.5	9.0	1066	1	ITAA_HUMAN
30	530	9.0	1048	1	ITAV_HUMAN
31	514.5	8.8	1044	1	ITAE_CHICK
32	511.5	8.7	1091	1	ITAE_MOUSE
33	505	8.6	1179	1	ITAV_MOUSE

RESULT 1

ALIGNMENTS

ID	ITAM_HUMAN	STANDARD	PRT	1152 AA
AC	P11215			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
DE	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	497	8.5	1396	1	ITA2_DROME	P12880 drosophila
35	491.5	8.4	1146	1	ITA1_DROME	Q24247 drosophila
36	490	8.3	1039	1	ITAB_HUMAN	P08514 homo sapien
37	489	8.3	1025	1	ITAB_HUMAN	P53708 homo sapien
38	488	8.3	1033	1	ITAB_MOUSE	O9qum0 mus musculus
39	486	8.3	1126	1	ITAM_MOUSE	P11578 cavia porce
40	478	8.1	1181	1	ITAM_HUMAN	Q13683 homo sapien
41	470	8.0	1106	1	ITA7_RAT	Q63258 rattus norv
42	466	7.9	1226	1	PAT2_CAEEL	P34446 caenorhabdi
43	444.5	7.6	1139	1	INA1_CAEEL	Q03600 caenorhabdi
44	419	7.1	1115	1	ITA3_DROME	O44386 drosophila
45	382	6.5	1000	1	ITA5_DROME	O9wlm8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=92144986; PubMed=1346576;
RX Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
conservation across species and homology to platelet IIB/IIIA.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18)." ;
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?" ;
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Saver R.W., Bryant G.L. Jr., Curry K.A.,
FA Fairbanks M.B., Finkel R.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Wildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment." ;
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface." ;
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide Cdl1b entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cdl1b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAU51960.1; -
DR PIR; A31108; RWHU1B.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1M1U; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWF_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

RESULT 3.

```
ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Last Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95
DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11C.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95.";
RL ENBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -1- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAMES=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81695; AAA59180.1; -
DR EMBL; Y00093; CAA68283.1; -
DR EMBL; M29165; -: NOT ANNOTATED CDS.
DR EMBL; M29487; AAA51620.1; ALT SEQ.
DR EMBL; M29482; AAA51620.1; JOINED.
DR EMBL; M29483; AAA51620.1; JOINED.
DR EMBL; M29484; AAA51620.1; JOINED.
DR EMBL; M29485; AAA51620.1; JOINED.
```

```
DR EMBL; M29486; AAA51620.1; JOINED.
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWF_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163
FT DOMAIN 20 1107
FT TRANSMEM 1108 1128
FT DOMAIN 1129 1163
FT REPEAT 34 87
FT REPEAT ? ?
FT DOMAIN 165 351
FT REPEAT ? ?
FT REPEAT 402 453
FT REPEAT 455 517
FT REPEAT 518 576
FT REPEAT 581 633
FT CA_BIND 466 474
FT CA_BIND 530 538
FT CA_BIND 593 601
FT SITE 1131 1135
FT DISULFID 69 76
FT DISULFID 108 126
FT DISULFID 655 712
FT DISULFID 771 777
FT DISULFID 848 863
FT DISULFID 998 1022
FT DISULFID 1027 1032
FT CARBOHYD 61 61
FT CARBOHYD 89 89
FT CARBOHYD 392 392
FT CARBOHYD 697 697
FT CARBOHYD 735 735
FT CARBOHYD 899 899
FT CARBOHYD 939 939
FT CARBOHYD 1050 1050
FT CONFLICT 490 490
FT CONFLICT 756 756
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62AA73 CRC64;
Query Match 58.6%; Score 3446; DB 1; Length 1163;
Best Local Similarity 60.5%; Pred. No. 1.5e-227;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAQPEI VAAHQSGSLYOCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFGSDSVQVANSVWVGAPQKITAAQTGLGYCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHTCSENTYVKGCLFLFGSNLRQOPK 120
DB 80 GLQVPEAVNMSLGLSLASTTSPQLACGPTVHHECGRNMVLTGLCFLLGPT--QLTOR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKKSKTLFSLMQVSEF 180
DB 138 LPVSRQECPCRBQDQIVFLIDGSGSISSRNFMFVRAVISQFORPSTQFSLMQFSNKF 197
```


FT	CARBOHYD	89	89	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	188	188	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	649	649	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	670	670	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	726	726	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	730	730	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	862	862	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	885	885	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	897	897	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	1060	1060	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. . .)	(POTENTIAL) .
FT	VARSPLIC	954	954	Q -> QGVHGLVEMOTSKOILCRPAGDAEHTVCAQGEPLCP	
FT				PWGVSEAFRDNTRAGPCR (in isoform 2).	
FT				/FTIG=VSP 002738.	
FT				R -> W (IN REF. 1 AND 2).	
FT				Y -> I (IN REF. 2).	
FT	CONFLICT	214	214		
FT	CONFLICT	660	660		
FT	STRAND	155	162		
FT	STRAND	164	164		
FT	TURN	165	166		
FT	HELIX	169	185		
FT	TURN	186	188		
FT	STRAND	191	198		
FT	STRAND	202	206		
FT	HELIX	208	213		
FT	HELIX	217	221		
FT	TURN	222	223		
FT	STRAND	229	229		
FT	HELIX	233	243		
FT	TURN	244	244		
FT	HELIX	247	249		
FT	TURN	250	250		
FT	STRAND	253	254		
FT	HELIX	274	276		
FT	TURN	277	278		
FT	STRAND	280	286		
FT	HELIX	288	290		
FT	TURN	293	297		
FT	TURN	298	299		
FT	HELIX	300	302		
FT	HELIX	307	310		
FT	STRAND	311	314		
FT	TURN	318	318		
FT	HELIX	319	328		
FT	TURN	329	330		
SQ	SEQUENCE	1170 AA;	128819 MW;	39A7AF92BF286FC0 CRC64;	
Query Match					
Best Local Similarity 34.2%; Score 1536.5; DB 1; Length 1170;					
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;					
QY	1	ENLDTENAMTFQ--ENARFGQSVVOLQGSRRVVVGAPQEIIVANORGSLYQCDYSTGSC	58		
Db	26	YNLDVVGARSFSPRAGRHFYGVYLVQV-GNGVTVGAPGE---GNSGSLYQCSGFGHCL	81		
QY	59	PIRLQVPEAVNMSLGLSLAATSPPLQLACGPTVHTQCSNTYVKGCLFLFGSNLR---	115		
Db	82	PVTLR-GSNVTSKYLGTATDPTDGSILACDPLSRTCDQNTYLSGLCYLFRQNLQGP	140		
QY	116	-QOPQFPALRCPOEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKXSKTLFSLM	174		
Db	141	LQGRPGFQECIKG---NVDLVFLFDGSMQLQDFQKILDFMKDVKMKLSNTSYQFAAV	196		
QY	175	OYSEEFRIHTFKFQNNPRSLIKPIITQLLGRTHATCLRVVPELFWITNGARKNAF	234		
Db	197	QESTSYKTFBFDSDYVYKRDQDALLKHVXHMLLTNTFGAINVATEVFREELGAPDAT	256		
QY	235	KILFLLTLDGKFGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRD	294		
Db	257	KVLIIITDGE--ATDSGNIDAAD-----IIRVIIGIKHFQKESQETLHKFPASKP	309		
QY	295	HVFQANNFALKTVQNLREKIFAIEGTQTGSSSFEHEMSQBGSAATISNGPLISTVG	354		

Db	310	FVKILDTPEKLDLFTLELQKKIYIEGTSKQDLTSFNMELSSSGISADLSRSHAVVGAVG	369		
QY	355	SYDMAGGVF-LYTSKEKSTFFINMTVDSMDMDAYLGAAAA-IILRRVQSLVGLGAPVQH	412		
Db	370	AKDMAGGFLDLKADLQDDTFIGNBPLTPEVRAGYLGTYVTWLPSPKQKTSLLASGAPRYQH	429		
QY	413	IGLVAMFR--QNTGMWESNANVKTQICAGYFCAGCSLDVDSNGSTDLVLICAPHYBOT	470		
Db	430	MGRVLLFOEPQGGHMSQVQTIHGTOIGSYFGGELCGVDVQDGETELLIGAPLYGEQ	489		
QY	471	RGQGVSVCPPLRGQRARWQCDAV--LYGEOGPMGRFGAALTVLGDVNGDKLTDVAIGAP	528		
Db	490	RGRGVFIY-----QRRQLGFVEVSELDQDGPVPLGRFGAITALTDINGDGLVDVAVGAP	544		
QY	529	GEEDNRGAVYLPHTGSGISPSHQSRIAGSKSLPRLQYFGQSLSGGQDLTMDGLVDLT	588		
Db	545	LEE--QGAVYIFNGRHG-GLSPQSPQRIEGTQVLSGIQWFGRSIHGVNDLEGDLAVAV	601		
QY	589	GAGQHVLLRSQPLRVKAIMFNPREVARNVFENDQVV-KGKAGEVRVCLHVQKSTR	647		
Db	602	GAESQMIVLSSRPVDMVTLSFSPAETPVHEVECSYSTSNMKKEGVNITICFOI-KSLY	660		
QY	648	DLREGOIQSVVTVYDLALDSGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQNPCL	707		
Db	661	POF-QGRIVANLTYTLQDGHTRRRGLFPGGRHELRRNIATV-TSMCTDIFSFPVPCV	718		
QY	708	EDVSPPIVLRNLFSL---VGTPLS--AFGN-----LRPVLAEDAORLFTALPFPKNCN	757		
Db	719	QDLISPINVSLNFWLWEEBEGTPRQRAQKQDIPPIRLPSLHSETWEI-----PFEKNCE	773		
QY	758	DNICODDLSITPSFMSLCLVVGPRFNVTVVRNDGEDSYRTQVTFPFPDLDSYRKVS	817		
Db	774	DKKCEANLRVSPSPARSRALRLTAPASLSVELSLNLEEDAYWVQLDHFPPGLSPRKVE	833		
QY	818	TLQORSQSRWLACES--ASSTEVSGALKSTSCSINHPIPFENSEVTNITFDVDSKAS	875		
Db	834	ML---KPHSQIPVSCCELPEESRLLSRAL---SCHVSSPIFKAGHSVALOMMFTLVNSS	887		
QY	876	LGNKLLKANVTSENN---MPRTNKTBFQLELPVKYAVYVVTSHGVSTKYNFTASEN	931		
Db	888	WGSVELHANVTNNEDSDLEDNSATTI---IPILYPINILIQOEDSTLVSVFPPKGP	944		
QY	932	TSRVNQHQYQV---SNLQORSIP-LSLVLPVRLNQTVWDRPQVTFSENLSSTCHTK-	986		
Db	945	KIQVKHMYQVRIQPSIHDHNIPTLEAVVGVQPPSEGIPTHQWSQVMEPPV--PCHYED	1002		
QY	987	RELPSHSD--FLAELRKAPVVCISIAVCORLQCDIPFGFQIEEFNATLKGNSLSPDWYIK	1043		
Db	1003	LERLPDAAEPCLPGALFRCPVV-----FROEILVQVIGTLELVGEIE	1044		
QY	1044	TSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTETKVEPEVENPLPLIVGSSVGGLL	1103		
Db	1045	AS-SMPSLCSLSISFNSKPHLYGSNASL-AQVVMKVVDVYVYEQMLYLVLSGIGLL	1102		
QY	1104	LLALITAAALYKLGFPKQYKQKMMSEG-GPPGAP	1136		
Db	1103	LLLLIFIVLYKGVFKRNLKEKMEAGRGVPGNP	1136		
RESULT 6					
ITAL MOUSE					
ID	ITAL MOUSE	STANDARD;	PRT;	1163 AA.	
AC	P24063;				
DT	01-MAR-1992 (Rel. 21, Created)				
DT	01-MAR-1992 (Rel. 21, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Integrin alpha-L precursor (Leukocyte adhesion glycoprotein IFA-1				
DE	alpha chain) (Leukocyte function associated molecule 1, alpha chain)				
DE	(CD11a).				
GN	ITGAL OR IFA-1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;
Query Match 19.3%; Score 1136.5; DB 1; Length 1167;
Best Local Similarity 28.6%; Pred. No. 1.6e-69;
Matches 353; Conservative 213; Mismatches 462; Indels 205; Gaps 43;
QY 1 FNLOTENA--WTFQENARGFGQSVVLOGSRVVVGAQOEIVAAHQGS-----LYOCDY 52
DB 20 FNMVDVWAWTALQFGAPAVLSSLLHLDPSN-----NOTCLLVARRSSNRNTAALYRCAL 74
QY 53 STGCEPIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHQTCSEYTYVGLCLFL 109
DB 75 SI-SPDEIACQ-PVEHICMPKGRYGQVTLVGNHNGVLVCIQVQAKFRSLNSELTGACSL 132
QY 110 FGSNLROQPKFPKALRG-----C-----PQE 131
DB 133 LTPNLDLQAQAYFSDLEGLDPGAHVDSGDYCRSGKSGTGBEEKSARRRRRTVEDEED 192
QY 132 DSDIAFLDGSIIIPDPRMKELVSTIMEQL--KSKTFLSLMOYSEEFRIHFTPEF 189
DB 193 GTEIAVLDSGSGSPDFQAKNFISTMMNFVEKFCFECFALVQVGAIVTQDFDQES 252
QY 190 QNNPNRSLIKPITQLLGRTHATGLRKVRLEFNITNGARKNAFKILFLTDGKFGDP 249
DB 253 RDINASLAKVSIQVQKEVTTASAMQHVLDNIPIPSGRSKKAKVMVLTDCGIFGDP 312
QY 250 LGYEDVIPEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQFQANFALKTVQ 309
DB 313 LNLTTVINSKMGVVRFAIGVGRFKNNTYRELKLIASDPKEAHTPKVTNYSALDGLL 372
QY 310 NQLEKPIAIGTQSGSSSEHEMSQEGFAAITSNGP--LLSTVGSVDWAGGVFLY--TS 367
DB 373 SKLQORIYHMEGT---VGDALQYLAQTGFGSAQILDKGQVLLGTVGAFNWSGGALLYSTQ 429
QY 368 KEKSTFINMT--RVDS--DMNDVAGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFRONTGM 425
DB 430 NGRGCFNLQTAKEDESRITQYSLGSLAVLHKAGCISVACAPRHLKAGVAFELKEDR- 488
QY 426 WESNA---NVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSVCLPLR 482
DB 489 -EEDAFVRRIEGEQMGVFGSVLCVPVDIDMGTTDFLLVAAPFVHIRGEEGRVYVQPE 547
QY 483 QARARWQCDVLYGEOGQWGRFGAALTIVLDVNGCDKLTDAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGLTNSRFGFAMAAGVDINDQKFTDVAIGAPLEFGAGGASVGS 606
QY 537 VYLFHGTSGSISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLL 596
DB 607 VYINGHSG-GLYDPSQIRASSVASGLHYFGMSVSGLDPNGDGLADIIVGSRDSAVV 665
QY 597 LRSQVPLVKAIEMFNPEARVARNFECNDQVVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMTFTF-----DALPMVFIGKM--DVNLCFEVDSSVVAEPGLREM 715
QY 654 QIQSVVTVYDLALDSRPHRAVFNETKSTRTQVGLGTQTC-----696
DB 716 FLNFTVDV-----TKQRQLQCEDSSGCSCLKRWNGSFLCBHFWLI 760
QY 697 ETLKQLPNCIEDPSPVLRNLSLVGTPLSAFNGLR-----PVLAEADQRLFTALF--P 750
DB 761 STEEL-----CEDCFSTNITIKTYE-----FOTSGRRDYNPTLI--DHYKEPSAIFOLP 809
QY 751 FEKNGCNDNICDDLSITFSFMSLDCLVVGGPRFNFVTVVRNDEGDSYRTQVTFPLD 810
DB 810 YEKDCKNVFCIAEIQLTIN--ISQELVVGTVKEVTMTNISLTNSGDSYMTNMAIYNPRN 868
QY 811 LSKYKSTLQNGORSQRSLRACASSSTVSALKSTCSINHPIPPENSEVTFNITDV 870
DB 869 LQFKKI-----QKVPSPDQVQDDPKPV---ASVLVNMCKIGHPII--KRSSVNVSVTWOL 918
QY 871 DSKASLGNKLLKANVTSENNMPRTNKTEFQLELPVKVAVYVMTVSHGVSTKYNFTASE 930

DB 919 EBSVFPNRTADITVTISNKSLSARETR---SLOPRHAFIATLSR--PSVMYMN--TSQ 971
QY 931 NTSRVNQHOYQVSNLQSRSLPISLFLVPLVRLNQTIVMDRPOVTFSENLSST-----CHT 985
DB 972 SPBDHKFEFNVHGENLFGAVFQLOICVPIKLODF-----QIVRVNLTKTODHTECTQ 1025
QY 986 KERLPSSHDSLAELRKAPVNVCSIAVCQRIQCDIPFGIQBEFNATLKNLSPDWYIKTS 1045
DB 1026 SQEPACGSDPVQHVKEHWSVVCAL-----TSNKENVTVVAABISVG 1065
QY 1046 HNFLLIVSTA-----EILFNDVSFTLLPGQAGFVRSQETKVEPF-----EVENPLPLIV 1095
DB 1066 HTKQLLRDVSLEPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEETRSPLII 1119
QY 1096 GSSVGGLLLALITAAALYKLGFPKROYKDMSE 1128
DB 1120 GSSIGULLVLVIIILFKCGFFKRYQQLNLE 1152
RESULT 8
ITAE HUMAN
ID ITAE_HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
DE ITCAE.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RP TISSUE=Leukemia, and Lymphocytes;
RC MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
[2]
RN REVISIONS TO 88-114.
RP Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 53-1179 FROM N.A.
RP TISSUE=Fetal kidney;
RC MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Aniketer Y., Dietrich N.L., Maduro V.V., McDowell G., Shetlerauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RT Genome Res. 10:165-173(2000).
[4]
RN MUTAGENESIS OF ASP-109 AND PHE-316.
RP MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
 CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L25851; AAB59359.2; -
 CC EMBL; AF168787; AAF43107.1; -
 CC PIR; A53213; A53213.
 CC HSSP; P11215; IABX.
 CC Genew; 604682; -
 CC MIM; 604682; -
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 3.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS0234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Polymorphism; Magnesium; Calcium.
 FT SIGNAL 1 18
 FT CHAIN 19 1179
 FT CHAIN 19 1179
 FT CHAIN 19 1177
 FT CHAIN 179 1177
 FT CHAIN 19 1124
 FT TRANSMEM 1125 1147
 FT DOMAIN 1148 1179
 FT DOMAIN 181 198
 FT REPEAT ? ?
 FT REPEAT ? ?
 FT REPEAT ? ?
 FT DOMAIN 145 199
 FT DOMAIN 200 391
 FT REPEAT 401 456
 FT REPEAT 457 506
 FT REPEAT 510 571
 FT REPEAT 573 638
 FT REPEAT 641 693
 FT CA_BIND 522 530
 FT CA_BIND 586 594
 FT CA_BIND 654 662
 FT SITE 1150 1154
 FT DISULFID 70 79
 FT DISULFID 126 159
 FT DISULFID 706 762
 FT DISULFID 823 829
 FT DISULFID 893 907
 FT DISULFID 1008 1033
 FT DISULFID 1041 1057
 FT CARBOHYD 49 49
 FT CARBOHYD 271 271
 FT CARBOHYD 321 321
 FT CARBOHYD 444 444
 FT CARBOHYD 726 726

FT CARBOHYD 782 782
 FT CARBOHYD 857 857
 FT CARBOHYD 934 934
 FT CARBOHYD 954 954
 FT CARBOHYD 1065 1065
 FT CARBOHYD 1096 1096
 FT VARIANT 360 360
 FT VARIANT 1041 1041
 FT MUTAGEN 208 208
 FT MUTAGEN 316 316
 FT CONFLICT 477 477
 FT CONFLICT 482 482
 FT CONFLICT 950 950
 FT CONFLICT 1019 1019
 FT SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;
 SQ
 Query Match 19.2%; Score 1128; DB 1; Length 1179;
 Best Local Similarity 28.7%; Pred. No. 6.3e-69;
 Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;
 QY 45 GSLYQCDYSTGS--CEPT-RLQVP-----VEAVNMSLGLSLAATTSPQLLACQPTVHQ 95
 DB 65 GELHRCSLVQDEILCHPVEHPVPKGRHGRVTVVRSHGVLCI-----QVLVRRP--HS 117
 QY 96 TCSENTYVKGCLFLFGSLNRQPOQ-----SDIAFLIDGSGSIIPHDFRMKELVSTIMEQL-- 119
 DB 118 LSSELT---GTCSLGLPDLRPOAQANFFDLENLLDPDARVDTGDCYSNKEGGEDDVNTA 174
 QY 120 KPPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRMKELVSTIMEQL-- 164
 DB 175 RQRALKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEE 234
 QY 165 KSKTFLPSLMQYSEEFRIHFTKEFQNNPNPSLKLPIQTLLGRTHATGLRKVYRELFN 224
 DB 235 KCFCNFALVQGGVIQTEFQDLRDSQDVMSLARVQNTQVGSVTKTASAMQHVLDISFT 294
 QY 225 ITNGARKNAFKILFLLTDGEKFGDPLGYEDVPELDREGVIRYVLGFGDARPEKSRQEL 284
 DB 295 SSHGRRKASKVMVLTGGIFEDPLNLTIVINSQMGQVERFAIGVGEFKAARTAREL 354
 QY 285 NTVASKPPRDHFVQANFEALKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 344
 DB 355 NLIASDPDETHAFKVTNYMALDGLSLKLYNIISMEGT---VGDALHYQLAQIGFSAQL 411
 QY 345 SNGP--LLSTVGSYDWAGGVLY--TSKEKSTFINMTVRVSDMND-----YLGVAIAIIRN 398
 DB 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAADAAADAAQAQYGLGVAVLHKT 471
 QY 399 RVQSLVLGAPRYOHIGLVAMFR--QNTGMWESNANY--KGTQIGAYFGASI--CSVDVDSNGST 456
 DB 472 CSLSVAGAPQYKHG--AVFELQKEGREASPLVLEGEQMGVSGFSELCPVDDIMDGGST 529
 QY 457 DLVLIGAPHYEQTRGGQVSVCPPLRGQARMQCDVILYGEQGPWGFAGAAITVLGDVN 516
 DB 530 DFLVVAAPFYHVHGEGRVYVYRLSE--QDGSFSLARILSGHPGFTNARFGFAMAMGDLS 588
 QY 517 GDKLTDVAIGAP-----GEEDNR--GAVYLFHGTSGSGISPSHSQIRIAGSKLSPLRQYFQ 570
 DB 589 QDKLTDVAIGAPLEGFGADGDGASFGSVIYNG--HWDGLSASPSQIRASTAVAPGLQYFGM 647
 QY 571 SILSGGQDLTMDGLVDLTVGAGQGHVLLLRSQPVLVRKATMEFNPVARNVFNCDQVVK 630
 DB 648 SMAGGFDISGGGLADITVGTGQAVFVRSPVRLKVSMAFTPSALP-----IGF 697
 QY 631 KEAGEVVRVCLHVQKSTRDLREGQIQSVVTVLDALDSGRPHSRVAFNETKSNTRRQTVL 690
 DB 698 NGVNVRLCFEI--SSVTTASESGLEALLNFTLDVDVGKQRRRLQCSVRSCLGCLREWS 756
 QY 691 GLTQTCETLKLQLPN-----CIEDPVSPIVLRNLSLVTGTPLSAFGNLRPLVAEDAQRLFT 746
 DB 757 SGSQLCEDL--LLMPTGEGELCEDCFSNASVKVSYQL--QTPEGQTDHPQILDTRYEPFAI 814


```
QY 107 CLEFGLNLRQPOKFFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKK 166
Db 120 CSDVSPFQVNSIAP--VQECSTQ-LDIVVLGDSNSIYPWD-----SVTAFLNDLLKR 171
QY 167 -----SKTFLSMLQYSEERIHTFKFQFNPNPRSLIKPIITQQLOR--THTATGLKVR 220
Db 172 MDIGPKQTQVIGVOYGYENTHLENLNKSYSTEELVAARKIVQGGRTMTALGTDITARK 231
QY 221 ELFNINIGARKNAKFIKILFULTDCEKFGDPLGVEDVIPELDREGVIRVYLGFGDAPR--- 276
Db 232 EAFTEARGARRGVKKVWVIVTDGESH-DNHLKVKVIOCEDENIQBFSAIIGSYNRGNL 290
QY 277 -SEKSRQELNIVASKPRPHRVFOANNFEALKTVQNOLREKIFAIEGTQTCSSSSPEHMS 335
Db 291 STEKPEVEIKSIASEPTEKHFFNVSDALAVITVKTGLERIFALEATAQSAASFEMWS 350
QY 336 QEGFSAIITNGPLLSVGSYDWAGGVLYTSKE-----KSTP-INNRVDSMDNDAYLG 389
Db 351 QTGFSAHYSQDWYMLGAVGAYDNGVTVMQASQIIIPRNTTFNVESTKKEPL-ASYLG 409
QY 390 YAAAIILNRVQSL-VLGAAPRYOHIGLVAMFRONTGWMESNANVKGTQIGAYFGASLCSV 448
Db 410 YTVNSATASGDVLYTAGPRYNTQVVIYRMEDGNIKILQTLSEQLGSGVFGSILTTT 469
QY 449 DVDSNGSTDVLVLGAPHY-----YEQTR-GGQVSVCPPLPRGQARWQCD 492
Db 470 DIDKSDNTDILLGAPWYMCETEKEQGVVYVALNQTREYQMSLEPIKQTCSSRQHS 529
QY 493 VLYGEGOPWG-RFGAALTVLGVNGDKLTDVAIGAPGEDNRGAYVLPHGTSGSGISPS 551
Db 530 CTTENKNEPCGARFGTAIAVAKDLNLDGFDNDIVIGAPLEDHGGAYVYIHG-SGKTIRKE 588
QY 552 HSQRIAGSKLSPLOYFGSLSGGQDITMDGLVDLTVGAQGHVLLLRSPVLVRKAIMFEF 611
Db 589 YAGRIPSGGDKTLKFGSIHGMDLNGDGLDVTIIGLGGAAFPWSDVAVVKTMPF 648
QY 612 NPREVARNFECDQVVKKEAG--EVVRHLVQ-KSTRDLREGQIQSVVTVYLDLDSG 668
Db 649 EPKNVNIQKNCH--MEGETVCINATVCFEVLKSKEDTIVEADLQ---YRVTLDSL 701
QY 669 RPHSRAVNET-----KNSRQTVGLGTQTCETLKLQPCIEDPVPVLRNLFSL 722
Db 702 ROISRSFFSQTBQKRVQNIITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNL 753
QY 723 VGTPLSFGNLRVLAEDAQRFTALFPPEKNGNDNICODLSITFSFMSLDCLVVGGP 782
Db 754 T-DPENG-----PVLDDSLPNSVHEIYIPFAKOCNGKEKCLSDLSLHVATTEKDLLIVRSQ 807
QY 783 RE-FNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACASSTVS 841
Db 808 NDKFNVLTVKNTKDSAYNTRTIVHYSPLNVFSGIEAIQK-----SCSN----- 853
QY 842 GALKSTCSINHPIFENBEVNTITFDVDSKASLGN-KLLLKANTVSENMPRTNKTBF 900
Db 854 ---HNITCKGYFPFLARGEMVTPEKILFQFNTSMENVTIYLSATSDSEBPETLSDNVV 910
QY 901 QLELPKYAVYVWVTVTSHGYSKTVLNTFASNTSRVMOHQYQVSN-----LGORS----- 949
Db 911 NISIPKYEVGLQFYS-SASEYHISIAAETVPEVINSTEDIGNEINIFYLRKSGSPM 969
QY 950 --LPISLVF-----LVVRLNQTIVDRPQVTFSENLSSTCTKE-----R 988
Db 970 PELKLSIFPNMTSNGYPVLYPTGLSS-----SENANCRPHIEDPFSINSKK 1018
QY 989 LPSHSDFLAELRKAPVNCVSIACQRIQCDIPFGIOE-----EFNA 1030
Db 1019 MTTSTD---HLKRGTLDCNCTKATITCNLTSSDISQVNVSLILWKPTFIKSYFSSNL 1075
QY 1031 TLKGNLSFDWYIKTSHNHLIVSTAELLFNDSVFTLLPGGAFVRVQTEKVEPFPVNP 1090
Db 1076 TIRGEL-----RSENASLVLSN-----QKRELAIQISKQGLPGR 1111
QY 1091 LPL--IVGSSVGGLLLALITAAALYKLGFFKROYKDMWSE 1128
```

```
Db 1112 VPLWVILLISAFAGLLLLMLLILALWKIGFFKRPLKKMEK 1151
RESULT 10
ID ITAH HUMAN STANDARD; PRT; 1189 AA.
AC Q9UKX5; Q9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
ID ITGA11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
RT human integrin alpha11 subunit (ITGA11).";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kusche-Gullberg M., Sejersten T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
RT integrin. A collagen-binding, i domain-containing, beta(1)-associated
RT integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-CAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF109681; AAF01258.1; -.
CC EMBL; AF137378; AAD51919.2; -.
CC EMBL; AL359064; CAB94392.1; -.
CC HSSP; P17301; IAOX.
CC Genew; HGNC:6136; ITGA11.
CC MIM; 604789; -.
CC GO; GO:0008305; C:integrin complex; TAS.
```



```
QY 995 FLAELKAPVNCIAVCORIQCDDIPFGIOBEFNATLKNLSFDWY-----IKTSHNHL 1049
Db 1051 --EDLRAPQLNHSNDVVSINCIRLVP-NQEIFHLLGNL---WLRSLKALKYKSMKI 1104
QY 1050 LIVSTAEILLFNDVSFTLLPQGAFAVSOTETKVERPEVEN-----PLPLIVGSSVGGILL 1103
Db 1105 MVNAALQORQH-SPF-----IFREEDPSRQIVFEISKQEDWQVPIIIVGSTLGGILL 1155
QY 1104 LIALITAAALYKLGFFK--RQYKDMSEGPPGAEP 1136
Db 1156 LLALLVLALWKLGFRRSARRRE-----PGLDLP 1183

RESULT 11
ID ITA2_BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
GN ITGA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94193647; PubMed=7511592;
RA Kamata T., Puzon W., Takada Y.;
RT Identification of putative ligand binding sites within I domain of
RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
RL J. Biol. Chem. 269:9659-9663(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ. IT
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL-outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L25886; AAB59255.1; -
DR FIR; I45914; I45914.
DR HSP; P17301; IAOK.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
```

```
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER 1
FT SIGNAL <1 18 POTENTIAL.
FT CHAIN 19 1170 INTEGRIN ALPHA-2.
FT DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1122 1143 POTENTIAL.
FT DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 92 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 177 367 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 423 475 FG-GAP 4.
FT REPEAT 477 538 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 604 656 FG-GAP 7.
FT CA_BIND 488 496 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 616 624 POTENTIAL.
FT SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1146 1150 GFFKR MOTIF.
FT DISULFID 72 81 BY SIMILARITY.
FT DISULFID 669 726 BY SIMILARITY.
FT DISULFID 778 784 BY SIMILARITY.
FT DISULFID 854 865 BY SIMILARITY.
FT DISULFID 1008 1039 BY SIMILARITY.
FT DISULFID 1044 1049 BY SIMILARITY.
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 580 588 G -> V.
FT VARIANT 588 588 R -> K.
FT VARIANT 725 725 R -> S.
SQ SEQUENCE 1170 AA; 128929 MW; ECEFF1C5F2448FB1 CRC64;
```

```
Query Match 18.2%; Score 1071; DB 1; Length 1170;
Best Local Similarity 27.5%; Pred. No. 5e-65;
Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

QY 1 FNLDTENAMTQ-ENARFGSGSVQL---QGSRVVVGAPQEIIVAAQNGSLYQC---DYST 54
Db 19 YNVGLPKAKIFGSPSSQFGYAVQQFINPKGNWLLVSGPWSGFPKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGLC 107
Db 79 TTCEKLNLOTSTMSNVTMTNMSLGLTLRNVTGCGFLTCGPLWAQCGSQYTTGVC 138
QY 108 FLFGSNLRQQPKPPEALURGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLK-- 165
Db 139 SDVSPDF-QLRTSPAPAVQTCB-SPIDVVVWCDESNIIVPWB--AVKNFLKFXVQGLDIG 194
QY 166 KSKTLFSLMVSSEPRHIFTKFEQNNPNPSLIKPIQLL-----GRTHATGLRKVRE 221
Db 195 PTKTMGLIQIYANNPRVVFVNLNTFKSKD---EMIKATSQTFQYGGDLTNTKAIQYARDT 251
QY 222 LPNITNGARKNAFKILFLTLTQCKPGDPLGVEDVPIELDREGVIRY---VLGF--GDAFR 276
Db 252 AYSTAAGRPGATKVMVVVTIDGESH-DGSKLKAVIDQCNKONILRFGTAVLYLRNALD 310
QY 277 SEKSROELNTVASKPPRDHVFQANNFEALKTQVQNLREKIFAIEGTQTGTGSSSFEHMSQ 336
Db 311 TKNLIKEIKATIASIPTEHRHFVNSDEADLLEKAGTIGQIFSGIEGTVGQ-GDNFQWMSQ 369
QY 337 EGFSAAIT--SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
```

```
Db 370 VGFAEYSPQNNILMLGAVGYDMGTVVQKTPHGHLLIFSQAQFEQILQDRNHSSYLGS 429
Qy 392 AAILLRNRVQSLVLCAPRYQHIQLVAMPRQNTGMNESANV-----KGTOIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQVILSVN-----ENGNTVVIQSGDQIGSVFGSVL 484
Qy 446 CSVDVDSNGSTDVLIGAPHYYEOTR--GGQVSVCPRLPRGORARQCDAVLYGQGPWG 503
Db 485 CAVDVNKDITDVLVGVPMYNDLKKEGRVYLTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RGAALTVLGVNGDKLTDVAIGAPGEDNRCAYLPHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIALSDINMGDFNDVIVGSPLENQNSGAVIYNGHEGM-ILRYSOKILGSDRAF 600
Qy 562 SPLQYFGQSGCGQDLTMDGLVLTGVAQGHVLLRQOPVLVRKVAIMEFNPREVARNVF 621
Db 601 SHLQYFGESLDGYGLNGDSITDVSAGFGQVQVLMQSDADVSVDASFPPKKI--TLL 658
Qy 622 ECNDQVVKGEAGEVRVLHVQKSTRDLRREGQIOSVVTYDALD-----SGRPHRAVFN 677
Db 659 NKNABE-----KLKLCF---SAKFRPTNQNNQVAIVYINITIDEDQFSSRVISGLFK 707
Qy 678 ETKNSTRROTQVLGTLQICE--TLKLQLPNCIEDPVPVLRLNFSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMIVSQAQRCSEYIIHQEPS-----DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVAEDAORLFTALFPPEKNGCNDNICODLSITF----SPMSLDCLVGVGPRFNVTV 789
Db 757 NFALAEYSTVKVFPFPHKCGDGVCGISDLVNLVQQLPATQOPQFFVIVNQNKLTFSV 816
Qy 790 TVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESAST-EVSGALKSTS 848
Db 817 QLKNNKESAYNTEIVVDFSENLF-----ASWMPVDGTEVTCQIASQKSVT 864
Qy 849 CSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKY 908
Db 865 CNVGPALKSKQVTFITNDFNLQ-NLQONQASISFRALSESQEEENMADNSVNLKSLLY 923
Qy 909 AVYVMTVSHGVSTKYLNFNTASENTRVMOHQVQVNLQOR-----SLPISLVFLV 958
Db 924 DAEIHT-RSTNINFEVSLDGNVASV-HSFE--DIGKPFIFSIXVTGSPVPSWA--- 976
Qy 959 PVRLNQTVIMDRPQVTFSEN-----LASTCHTK-----RLPHSHDFLAE- 998
Db 977 -----SVIIHIPQVTKDNPLMYLTGVHTDQAGDISCEAENPLKIGTSSSVFSKSEN 1030
Qy 999 LRKAPVNCISAVCQBIQCDIPFGIQEFPNATLKNLSFDWYIKTSHNHLIVSTAET- 1057
Db 1031 FRHIKELNCRATASCSNIMCWLRLQVKGIFYLNVSTRIWNGTFAASTFOTVQLTAAAEID 1090
Qy 1058 LFNDSVFTL-----LPGQAFVRSQTEKVPFE-VPNPLPLVIGSSVGGLLLALITA 1110
Db 1091 TYNPQIVYIEENTVTP-----LTIKPFHEKVEVPTGTVIGSVIAGIALLLALVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFKKRYEKM 1155

RESULT 12
ID ITA2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

FT	CA_BIND	560	568	POTENTIAL.	
FT	CA_BIND	624	632	POTENTIAL.	
FT	SITE	480	482	CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE	1154	1158	GFPR MOTIF.	
FT	DISULFID	80	89	BY SIMILARITY.	
FT	DISULFID	677	734	BY SIMILARITY.	
FT	DISULFID	786	792	BY SIMILARITY.	
FT	DISULFID	862	873	BY SIMILARITY.	
FT	DISULFID	1016	1047	BY SIMILARITY.	
FT	DISULFID	1052	1057	BY SIMILARITY.	
FT	CARBOHYD	102	102	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	109	109	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	429	429	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	472	472	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1054	1054	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1078	1078	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	1178	1178	AA; 128926 MW; IF194B9C0240F465 CRC64;	
Query Match					18.0%; Score 1057; DB 1; Length 1178;
Best Local Similarity					27.8%; Pred. No. 4.5e-64;
Matches					341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;
Qy	1	FNLDTENAMTQ--ENARFGQSVVQL---	QGSVVVGAQOEIIVAANQSGSLYQC--DYST 54		
Db	27	YNVGLPGAKIFSGSSQFGVSVQQLTNPQGNMLLVGSPWSPFENRMDGVYKCPVDLPT 86			
Qy	55	GSCPIRLO-----VPVEAVMSGLSLAATSPPOLACGTFVHQTSENYYKGLC 107			
Db	87	ATCEKUNLQNASISNVTETKNNLSGLTLTRNPGTGGFLTCPLMAHOCGNQYATGIC 146			
Qy	108	FLFGSLNRQOPQ---KPEALRGCPQSDIAFLIDGSGSIIIPHDFFRMKELVSTIMEQL 164			
Db	147	-----SDVSPDFQELTSPVAVQCPSL--VDVVVVCDESNIYP--WEAVKNFLVKPVTGL 199			
Qy	165	K--KSKTFLSIMOYSEBRFHFTKFEQNNPNRSLIKPIITQLLG-RTHPTATGLRKVVRE 221			
Db	200	DIGPKTKTQVALIOVANEPRIFNLNDPTEKEDVMQATSETROHGGDLTNTFRAIEFARDY 259			
Qy	222	LFNITGARNAKFILLLTDGKFGDPLGYEDVPELDEGVIRY---VLGF--GDAFR 276			
Db	260	AYSQTSGRGPATKVMVVVVDGESH--DGSKLKTVIOQCNDDEILRFGIAGVLGYLNALD 318			
Qy	277	SEKSRQLNTVASKPPRDHVFQANNFEALTKVQNLREKIFAJEGTGTGSSSSFEHMSQ 336			
Db	319	TKNLKIKKAIATPTETRYFENVADAEALKEAGTLGEQIFSIIEGTVOG--GDNFQWEMAQ 377			
Qy	337	EGFSA--AITSNGPLLTGVSYDWAGGVFLYTSKEKSTFINMT--RYVDSMN--DAYLGYA 391			
Db	378	VGFSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVLQDRNHSSFLGYS 437			
Qy	392	AAILNRVOSLVGAPRYOHIGLVANFRONTGWESNANV----KGTQIGAVFGASLCS 447			
Db	438	VAAISTBDGVHFVAGAPRANTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494			
Qy	448	VDVDSNGSTDLVLIGAPHYEQTR--GGQVSVCLPRGQARWQCCAVLYGEOGQPWGRF 505			
Db	495	VDVDKOTITDVLVGATTPNDLUKKEGKFTLFTITKGIUNHQH---FLEGPGTGNARF 551			
Qy	506	GAALTVLGDVNGDKLTVAIGAPGEEDNRGAVLYFHTGTSGSGISPSHQSRIAGSKLSR- 564			
Db	552	GSAIALSDINMDGFNDVIVGSPVENENSGAVIYNGHQT--IRTKYSQKILSGNGAFRR 610			
Qy	565	-LQYFGSLSGGQDLTDWGLVDLTVGAGQHVLLRSQPVLRVKAIMEFNPVARNVFE 623			
Db	611	HLQFFGRSLDGYGLNGDSITDVSIGALGQVQLWSQSIADVAIEALFTP----- 660			
Qy	624	NDQVVKGEAGEVRVCLHVQKSTDRLRREGQIQSVVTVYDLALD----SGRPHSRAVENET 679			
Db	661	-DKITLUNKAKITLKCFAERFPAGQNNQV--AILFNMTLADAGHSSRVTSRGVFRN 717			

Qy	680	KNSTRQTOVLGLTQTCET--LKLQLPNCIEDPVSIVLRNLPSPVLGTVPLSFGNRPVL 737			
Db	718	SERFLQKNVAVNEVQKSEHHSIQKPS---DVVNPDLRLVDISLENPGTS-----PAL 768			
Qy	738	AEDAQLFTALPFPEKNCNDNICODDLSI-----TFSPMSLDCLVVGCPREFNVT 788			
Db	769	EAYSETVKVFSIPFYKECGSDGICISDLILDVQQQLPAIQOTQSF-----IVSNQNKRLTFS 823			
Qy	789	VTVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACBSAST-EVSGALKST 847			
Db	824	VILKNRGESAYNTVLAEPSENLF-----ASFSPVDGTEVTCVSGSQKSV 871			
Qy	848	SCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKTE--FOLELP 905			
Db	872	TCDVGYPAALKSEQVTFETTFNFDENLQ-NLQNAAINFOAFSESQ--ETNKADNSVSLTIP 928			
Qy	906	VKYAVVMVYSTHGVSTKYLNFTASENTRVMQHQVQVSNLQOR-----SLPISLV 955			
Db	929	LIYDAELHUT-RSTWINFYEISDENAPSVIK---SVEDIGPKFIFSLKVTAGSADPVMA 984			
Qy	956	FLV-----PVRLNQTVIWRPQVTF-SENLS 980			
Db	985	LVTIHIPQVTKKPNLLYLGTQDQAGDISCTAEINPLKPHTA----PSVSFKNENR 1040			
Qy	981	STCHTKERLPSSHDFLAELRKAPVNCISAVCORIQCDIPFGIOEEFNATLKNLSFDW 1040			
Db	1041	---HTKE-----LDCRTTSCNITCMLKDLHMAEYFINVTTRVWNR 1080			
Qy	1041	YIKTSHNHLIVSTAEILFENDSVFTLLPGQAFVRSQETKVEPPEFVNPPLPLVGVSSVG 1100			
Db	1081	FAASTPQTVQLTAABEIDTHNPQLFVIBENAVTIPIMMKTEKAEVPT--GVIIGSIIA 1138			
Qy	1101	GLLLALITAAALYKLGFFKRYQKDM 1125			
Db	1139	GILLAMTAGLWKLGFKRYQKDM 1163			
RESULT 13					
ID	IT2A2 HUMAN	STANDARD;	PRT;	1181	AA.
AC	P17301;				
DT	01-AUG-1990	(Rel. 15, Created)			
DT	01-AUG-1990	(Rel. 15, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)				
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).				
GN	ITGA2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]_SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.				
RP	TISSUE=Endothelial cells;				
RC	MEDLINE=89308879; PubMed=2545729;				
RX	Takada Y., Hemler M.E.;				
RA	"The primary structure of the VLA-2/collagen receptor alpha 2 subunit				
RT	(platelet GPIa): homology to other integrins and the presence of a				
RT	possible collagen-binding domain.";				
RL	J. Cell Biol. 109:397-407(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,				
RA	Rajkumar N., Toth E.J., Yi O., Nickerson D.A.;				
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.				
RX	MEDLINE=98019223; PubMed=9353312;				
RA	Emsley J., King S.L., Bergelson J.M., Liddington R.C.;				
RT	"Crystal structure of the I domain from integrin alpha2beta1.";				
RL	J. Biol. Chem. 272:28512-28517(1997).				
RN	[4]				
RP	VARIANT HPA-5 (BR).				

RA MEDLINE=94043762; PubMed=7901236;
 RA Santoso S., Kalb R., Walke M., Kiefel V., Mueller-Eckhardt C.,
 RA Newman P.J.;
 RT "The human platelet allantoicins Br(a) and Brb are associated with a
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
 RT alpha 2).";
 RL J. Clin. Invest. 92:2427-2432(1993).
 RN [5]
 RP VARIANT GLU-534.
 RX MEDLINE=20206009; PubMed=10744142;
 RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
 RT gene polymorphism on coronary artery disease and acute myocardial
 RT infarction";
 RL Thromb. Haemost. 83:392-396(2000).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
 CC associates with beta-1. Interacts with HPSS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
 CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
 CC THROMBOCYTOPENIA (NAIT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
 CC ROLE IN CORONARY ARTERY DISEASE (CAD).
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17033; CAA34894.1;
 DR EMBL; AF512556; AAM34795.1;
 DR PIR; A33998; A33998.
 DR PDB; 1AOX; 25-NOV-98.
 DR PDB; 1DZ1; 02-AUG-01.
 DR Genew; HGNC:6137; ITGA2.
 DR MIM; 192974;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0005518; F:collagen binding activity; TAS.
 DR GO; GO:0007596; P:blood coagulation; TAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRIN.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
 FT SIGNAL 30 1181 INTEGRIN ALPHA-2
 FT CHAIN 30 1132 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1133 1154 POTENTIAL.
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1155 1161 INTERACTION WITH HPSS.
 FT REPEAT 45 103 FG-GAP 1.
 FT REPEAT 2 FG-GAP 2.
 FT DOMAIN 188 378 VWFA.
 FT REPEAT 378 433 FG-GAP 3.
 FT REPEAT 434 486 FG-GAP 4.
 FT REPEAT 488 549 FG-GAP 5.
 FT REPEAT 551 610 FG-GAP 6.
 FT REPEAT 615 667 FG-GAP 7.
 FT CA_BIND 499 507 POTENTIAL.
 FT CA_BIND 563 571 POTENTIAL.
 FT SITE 1157 1161 GFFKR MOTIF.
 FT DISULFID 83 92 BY SIMILARITY.
 FT DISULFID 680 737 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT DISULFID 865 876 BY SIMILARITY.
 FT DISULFID 1019 1050 BY SIMILARITY.
 FT DISULFID 1055 1060 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 534 K -> E (IN ALLOANTIGEN HPA-5B;
 FT dbSNP:1801106).
 FT /FTid=VAR_003977.
 FT TURN 170 171
 FT STRAND 173 180
 FT TURN 183 184
 FT HELIX 188 199
 FT TURN 200 201
 FT STRAND 204 204
 FT TURN 206 207
 FT STRAND 209 216
 FT TURN 220 224
 FT STRAND 226 228
 FT TURN 232 240
 FT TURN 241 241
 FT HELIX 252 262
 FT TURN 263 264
 FT HELIX 266 268
 FT TURN 269 269
 FT STRAND 275 282
 FT HELIX 289 291
 FT HELIX 292 301
 FT TURN 302 303
 FT STRAND 304 311
 FT HELIX 313 317
 FT TURN 318 319
 FT HELIX 323 330
 FT TURN 331 332
 FT HELIX 337 340
 FT STRAND 341 344
 FT HELIX 347 353
 FT HELIX 354 362
 FT TURN 363 363
 SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
 Query Match 17.9%; Score 1054; DB 1; Length 1181;
 Best Local Similarity 26.8%; Pred. No. 7.3e-64;
 Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

ITAG_HUMAN
 ID ITAG_HUMAN STANDARD; PRT; 1167 AA.
 AC 075578; Q9UHZ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-10 precursor.
 GN ITGA10.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 SEQUENCE FROM N.A.
 RC TISSUE=Articular chondrocytes;
 RX MEDLINE=98352078; PubMed=9685391;
 RA Camper L., Hellman U., Lundgren-Akerlund E.;
 RT "Isolation, cloning, and sequence analysis of the integrin subunit
 alpha10, a beta1-associated collagen binding integrin expressed on
 chondrocytes.";
 RL J. Biol. Chem. 273:20383-20389(1998).
 RN [2]
 SEQUENCE FROM N.A.
 RC TISSUE=Endothelial cells, and Heart;
 RX MEDLINE=20169197; PubMed=10702680;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
 Wang S.-X., Langley R., Kriessaneen G.W.;
 RT "The integrin alpha10 subunit: expression pattern, partial gene
 structure, and chromosomal localization.";
 RL Cytogenet. Cell Genet. 87:238-244(1999).
 CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN
 CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF074015; AAC31952.1; -;
 CC EMBL; AF112345; AAF21944.1; -;
 CC EMBL; AF172723; AAF61638.1; -;
 CC HSP; P17301; IAOX.
 CC Genew; HGNC:6135; ITGA10.
 CC MIM; 604042; -;
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding activity; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR020335; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 4.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 CC PROSITE; PS00234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 1167
 FT DOMAIN 23 1122
 FT TRANSMEM 1123 1145
 FT DOMAIN 1146 1167
 FT REPEAT 38 97
 FT REPEAT ?
 FT DOMAIN 167 350
 FT REPEAT 365 427
 FT REPEAT 428 482
 FT REPEAT 483 545
 FT REPEAT 546 605
 FT REPEAT 608 660
 FT DOMAIN 1134 1140
 FT CA_BIND 494 502
 FT CA_BIND 558 566
 FT CA_BIND 620 628
 FT DISULFID 76 86
 FT DISULFID 666 675
 FT DISULFID 681 736
 FT DISULFID 789 795
 FT CARBOHYD 98 98
 FT CARBOHYD 234 234
 FT CARBOHYD 336 336
 FT CARBOHYD 364 364
 FT CARBOHYD 733 733
 FT CARBOHYD 763 763
 FT CARBOHYD 839 839
 FT CARBOHYD 921 921
 FT CARBOHYD 1011 1011
 FT CARBOHYD 1018 1018
 FT CARBOHYD 1039 1039
 FT CONFLICT 844 844
 FT CONFLICT 909 909
 FT CONFLICT 926 926
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;
 Query Match 17.7%; Score 1042.5; DB 1; Length 1167;
 Best Local Similarity 28.5%; Pred. No. 4.4e-63;
 Matches 348; Conservative 201; Mismatches 495; Indels 177; Gaps 44;
 QY 1 FNLDTENAMTFOENARG-FGQSVVLOQSGR-----VVGAPQEIIVANQROSLVOC----- 50
 DB 23 FNLDHHPRLFGPPPEAFSGYSLVQVGGQGRWMLVGAPWDGPGSDRRGDDVRCVGGAH 82
 QY 51 -----DYSTG-SCEPIRLOVPVEAVNMVLSLAATTSPPQLLAGCTVHTQCS 99
 DB 83 NAPCAKHLGDYQLGNSSHP-----AVNMHLGMSLETDGCGGFMACPLWRACGS 134
 QY 100 NTYVKGCLFGLFGSNLRQPKPEALRGCPQSDSDIAFLIDGSGSIIPHD-----FRMK 154
 DB 135 SVFSSGICARVDASFQPGQSLAPTQR-CPTY-MDVIVLDGNSIYPHSEVQTFRLRLV 192
 QY 155 ELVSTIMEQLKSKTFLSLMQVSEBFRIFHTPKFQNNPNRSLIKPITQLLGR-THTAT 213
 DB 193 GKLFIDPEIQ-----VGLVQYGSPVHESLGDFTKEEVVRAAKNLSRRGRETAKTAQ 247
 QY 214 GLRKVVRELFNITNGARKNAFKILLLTDEKEF-GDPLGYEDVIBELDEGVIRY-VLGF 271
 DB 248 AIMVACTGFGSQSHGRPEEARLLVVVDGSHDGEEL--PAALKACEAGRVTYRGIAVL 305
 QY 272 GDAPRSEKS-----ROELNVTASKPRDRHVFOANNFEALNTQONQRLREKIFAJEGTQGS 327
 DB 306 GHYLRQRDPSPSFLREIRTIASDPDERFFNVTDAAALTDIVDALGDRIFGEGSHAENE 365
 QY 328 SSFEHEMSQEGFSAAITNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDS-----D 382
 DB 366 SSFGLMSQIGFSTHRLKDGILFGMVAGYDMGGSVLWLEGGHRLFPFRMALEDEFPALQ 425
 QY 383 MNDAYLGYA-AAIILNRVQSLVLCAPRYOHLGLVAMER-QNTGHWESNANVKGQIGAY 440
 DB 426 NHAAYLYGVSVMLLRGRRRLFLSGAPRFRHGRKVIAFQKKDGAIVRAQSLQEQIGSY 485
 QY 441 FGASLCSVDVDSNGSTDLVLICAPHYY--EQTGRGQVSVCPPLRQORARWQCDVLYGQ 498

Db 486 FGSELCPDTRDGTDDVLLVAAPMFLGPONKETGRVYVYLV--GQSSLLTLOQTLOPEP 543
Qy 499 QPWGRFCAALTVLGDVNGDKLTVAICAPGEEDNRGAVYLFHGTSGSGISPSHSORLAIAG 558
Db 544 PQD-ARFGFANGALPDLNODGFADVAVGAPLEDGHQGALYLYHGTQ-SGVRPHPAQRIAA 601
Qy 559 SKLSPRLQYFGOSLSGGDLTMDGLVDLTVGAQGHVLLLRSPVLRVKAIEMFNPREVAR 618
Db 602 ASMPHALSYFGRSDVGRDLDDGDDLDVAVGAQGAAILLSRPVHLTPSLEVTPOAISV 661
Qy 619 NVFECNDQVVKKEAG--EVRVCLHVQKSTRDRREGQIQSVVYDYDALDSGRPHRAVF 676
Db 662 VORDCRR--RGQEAUCLTAALCEQVTSRTFCRWDH--QFYMRFTASLDEWTAGARAAF 715
Qy 677 NET--KNSTRQTVLGTQTCETLKLQLPNCIEDPVSPIVLRNLSVLTGTPLSAFGNLR 734
Db 716 DSGQRLSPRLRLSVG-NVTCEQLHFHVLD-TSDYLRPVALTVTFALDNTTKFG----- 768
Qy 735 PVLAEDAQRLLFTALPPFEKNCNDNICODDLSITPSPMSLDC-----LVVGGPREFN 786
Db 769 PVLNEGSPSTISQKLVFPFSGDCGPNCEVTDVLQ---VNMDIRGRKAPFVVRGRRKVL 825
Qy 787 VTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEVSGALKS 846
Db 826 VSTTLNRKENAYNTSLIIFSRNL--HLASLTPQR-ESPIKVECAAPSA-----HA 874
Qy 847 TSCSINHPIFPENSEVTNITFDVDSKASLG---NKLL-----LKAVTSENMPRTNKT 898
Db 875 RLCSVGHVPVFTGAKVTELEFEFSCSLLSQVFGKLTASSDSLSERNGTLOENTAQT--- 931
Qy 899 EFQLELPVKYAVYVTVTSHGVSTKYLNFTASNTSRVMOHV-----QVS 943
Db 932 -----SAYIOYEPH-----LFFSSSTLHRYEYHPYGTLPVCGPGEFKTTLVQ 975
Qy 944 NLG---QRSPLISLVLP-----VRLNQTVIMDRPQVTFSENLSSTCHTKERLPS 991
Db 976 NLGCYVWSGLIISA--LLPAVAHGGNYFLSLSOVI-----TNNASCIVQNLTPEPP 1024
Qy 992 HSDFLAELRKAPVNCSTAVCQRIQCDIPFFGIQEEFNATLKGNLSPDWYIKTSHNLLI 1051
Db 1025 PPVHPEELQHTNRLNGSNTQCVVRCHLQGLAKGTEVSVGLRLVHNEFFFRAPKPSLTV 1084
Qy 1052 VSTABILFNDVSFTLLPGQAFVRQVETKVEPFEVNPPLPLIVGSSVGGLLLLALITAA 1111
Db 1085 VSTFELCTEESVLOLTEASRWSLSLEV-VQTRPILISLWILIGSVLGGLLLLALLVFC 1143
Qy 1112 LYKLGFF-----KQYK 1123
Db 1144 LWKLGFFAHKKIPEBEKREEK 1164

Search completed: November 25, 2003, 14:17:26
Job time : 14.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 / Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFCQ.....FKRQYKDMSEGGPGCAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPTREMBL 23.:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4358	74.1	1151	Q9J130	Q9J130 rattus norv
2	3878.5	66.0	1036	Q8CA73	Q8CA73 mus musculus
3	3786.5	64.4	920	Q28984	Q28984 sus scrofa
4	3457	58.8	1169	Q81VAG	Q81VAG homo sapien
5	3297.5	56.1	1169	Q9QXH4	Q9QXH4 mus musculus
6	3225.5	54.9	1161	Q9QYV7	Q9QYV7 rattus norv
7	1512.5	25.7	1161	Q9QTV4	Q9QTV4 mus musculus
8	1502	25.5	1160	Q9R200	Q9R200 mus musculus
9	1382	23.5	1196	Q98TF1	Q98TF1 cyprinus ca
10	1341.5	22.8	1086	Q96HB1	Q96HB1 homo sapien
11	1329.5	22.6	1187	Q98TF0	Q98TF0 cyprinus ca
12	1264	21.5	927	Q8H2V0	Q8H2V0 bos taurus
13	1142.5	19.4	1167	Q88340	Q88340 rattus norv
14	1097	18.7	1167	Q88341	Q88341 rattus norv
15	1034	17.6	1171	Q42094	Q42094 gallus gall
16	1020	17.3	1038	Q8BS01	Q8BS01 mus musculus

17	993	16.9	895	11	Q9WUF8	Q9WUF8 mus sp. it9
18	986.5	16.8	1160	6	Q8MKF4	Q8MKF4 felis silve
19	865	14.7	348	4	Q8TES5	Q8TES5 homo sapien
20	837	14.2	1332	5	Q9BPQ8	Q9BPQ8 halocynthia
21	784	13.3	205	11	Q63001	Q63001 rattus norv
22	753.5	12.8	780	13	Q06271	Q06271 xenopus lae
23	738	12.6	823	4	Q8W118	Q8W118 homo sapien
24	686.5	11.7	823	11	Q8CE84	Q8CE84 mus musculus
25	672	11.4	1032	11	Q61989	Q61989 mus musculus
26	643	10.9	1036	11	Q91YD5	Q91YD5 mus musculus
27	639.5	10.9	1033	6	Q9BGU3	Q9BGU3 bos taurus
28	628.5	10.7	257	11	Q8C270	Q8C270 mus musculus
29	619.5	10.5	1041	5	Q9UB90	Q9UB90 lytechinus
30	614.5	10.5	1041	5	Q76378	Q76378 lytechinus
31	579.5	9.9	1054	5	Q9U6S1	Q9U6S1 strongyloce
32	579	9.8	1034	13	Q98TT7	Q98TT7 gallus gall
33	546	9.3	1033	13	Q42598	Q42598 xenopus lae
34	534	9.1	1016	13	Q91779	Q91779 xenopus lae
35	530	9.0	1047	6	Q9MZD6	Q9MZD6 bos taurus
36	529	9.0	974	11	Q924W2	Q924W2 rattus norv
37	529	9.0	1007	6	Q9GK48	Q9GK48 bos taurus
38	528	9.0	1073	11	Q8CC06	Q8CC06 mus musculus
39	513.5	8.7	1034	6	Q9TUN4	Q9TUN4 oryctolagus
40	510.5	8.7	1036	6	Q9TUN6	Q9TUN6 sus scrofa
41	507	8.6	1012	11	Q70304	Q70304 mus musculus
42	491.5	8.4	1049	5	Q8SY51	Q8SY51 drosophila
43	475	8.1	833	5	Q9BPQ7	Q9BPQ7 halocynthia
44	471.5	8.0	604	11	Q8BQ25	Q8BQ25 mus musculus
45	470.5	8.0	1036	6	Q9TU44	Q9TU44 canis famil

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT: 1151 AA.
 ID Q9J130
 AC Q9J130
 DT 01-OCT-2000 (TREMREL. 15, Created)
 DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
 DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr., Zerrila K. Jr.;
 RT "Cloning of the rat CD11b cDNA sequence."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268593; AAF81280.1;
 DR HSSP; P11215; 1BHQ
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWEA; 1.
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.1%; Score 4358; DB 11; Length 1151;
 Best Local Similarity 72.3%; Pred. No. 1.8e-314;
 Matches 822; Conservative 151; Mismatches 162; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGFCQSVVVOIGSRVVVVGAPQEIIVAAHQGSLYOCDYSTGSCPEI 60

DB 17 FNLDTENAMTFQENARGFCQSVIQLGETRVVAAPOEVKAVNQTGALYQCYDSTNRCDPI 76

```
Qy 61 RLQVPPAVNMSLGLSLAATTSPPQLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
Db 77 PLQVPPAVNMSLGLSLAATTVPPQLACGPTVHOTCKENTYVNGCLCYLFGSNLLRPQQ 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQSDSIAFLIDGSGSINTIDFQKMEFVSTVMDQFQKSKTILFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLIKIPITOLLGRTHATGLRKVRELFINITGARKNAKILFL 240
Db 197 RTHFTFDFKRNPPKSHVRPIQLNGRTKTASGIRKVRRELFOKINGADNAAKILW 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVYVLGDFADFRSEKSRQELNTVASKPPRPHVFOAN 300
Db 257 TDGKFGDPLGYEDVPEAEAGIIRVYVGNVAFHFKPQSRRELDTTASKPAGHRVFOVD 316
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWA 360
Db 317 NFEALNTIRNQLQEKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWA 376
Qy 361 GVFLYTSKEKSTFFINMTVRDSDMDNDAYLGYAAAILLRNVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFINTTRIDSDMDNDAYLGYASAVISRRNVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYETOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYETOTRGQGVSCPL 496
Qy 481 PRGORARWQCDAVLYGEGQPMWRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDAVLYGEGQPMWRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDITMDGLVDLTIVGAQCHVLLRSQ 600
Db 556 HGASVASISTSPHSQRIAGARESPGLQYFGQSLSGQDITMDGLVDLTIVGAQCHVLLRSQ 615
Qy 601 PVLKVKATMEPNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660
Db 616 PVLKVKATMEPNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 675
Qy 661 YDLALDGRPHSRVAFNFTKSTRQOTVGLTQTCETLKLQPCNICBDPVSPIVLRINF 720
Db 676 YDLALDGRPHSRVAFNFTKSTRQOTVGLTQTCETLKLQPCNICBDPVSPIVLRINF 735
Qy 721 SLVGTPLSAFNLRLPVLAEADQRLFTALFPPEKNCNGNDNICODDLSITFSPMSDCLVVG 780
Db 736 TLVGELPSSRDLRPLVLAERQRIFTAMPFPEKNCNGNDNICODDLSITFSPMSDCLVVG 795
Qy 781 GPREFNVTVRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 796 DSRDPSVTVLRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRORSWRLACESASSTEV 854
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 QGVLKSTIWDINHPIFPANSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 914
Qy 901 QLELPVKAVYVWVTSYKYNFTASENTSRVQHQYQVNSLQGRSLPISLVFLVPV 960
Db 915 QLELPVKAVYVWVTSYKYNFTASENTSRVQHQYQVNSLQGRSLPISLVFLVPV 974
Qy 961 RLNQTIVWDRPOVTFSENLSSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORTQCDIP 1020
Db 975 QINKVTIWDPPQVTFSENLSSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORTQCDIP 1034
Qy 1021 PFGIQEENATLKNLSFDWYIKTSHNHLIIIVSTAEIIFNDSVFTLLPGQAGFVRSQTE 1080
Db 1035 SPNSKEIENVTLQGNLLFDWYIKTSHNHLIIIVSTAEIIFNDSVFTLLPGQAGFVRSQTE 1094
Qy 1081 KVEPEVNPPLIVGVSSVGGILLALITAAALYKLGFFKRYQKMDMSGGPPGAEPPQ 1137
Db 1095 KVEPYTVHNPVPLIVGVSSVGGILLALITAGLYKLGFFKRYQKMDMSGGPPGAEPPQ 1151
```

```
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
ID Q8CA73
AC Q8CA73;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK039444; BAC30350.1; --
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E1B8E77 CRC64;

Query Match 66.0%; Score 3878.5; DB 11; Length 1036;
Best Local Similarity 65.6%; Pred. No. 6.4e-279;
Matches 747; Conservative 126; Mismatches 146; Indels 119; Gaps 2;

Qy 1 FNLDTENAMTQBNARGFGQSVVOLQGSRRVVGAPQIIVANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENAKFGQSVVOLQGSTVVAAPQBAKAVNQTAGLYQCDYSTSRCHPI 76
Qy 61 RLQVPPAVNMSLGLSLAATTSPPQLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
Db 77 PLQVPPAVNMSLGLSLAATTVPPQLACGPTVHOTCKENTYVNGCLCYLFGSNLLRPQQ 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSINTIDFQKMEFVSTVMDQFQKSKTILFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLIKIPITOLLGRTHATGLRKVRELFINITGARKNAKILFL 240
Db 197 RIHFTFDFKRNPPKSHVRPIQLNGRTKTASGIRKVRRELFOKINGADNAAKILW 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVYVLGDFADFRSEKSRQELNTVASKPPRPHVFOAN 300
Db 257 TDGKFGDPLGYEDVPEAEAGIIRVYVGNVAFHFKPQSRRELDTTASKPAGHRVFOVD 316
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWA 360
Db 317 NFEALNTIRNQLQEKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWA 376
Qy 361 GVFLYTSKEKSTFFINMTVRDSDMDNDAYLGYAAAILLRNVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGYASAVISRRNVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYETOTRGQGVSCPL 480
Db 437 ENFTGPHSTSIK----- 450
Qy 481 PRGORARWQCDAVLYGEGQPMWRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDITMDGLVDLTIVGAQCHVLLRSQ 600
Db 451 -----SQRIIGAHFSPGLQYFGQSLSGQDITMDGLVDLTIVGAQCHVLLRSQ 498
Qy 601 PVLKVKATMEPNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660
Db 499 PVLKVKATMEPNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 558
```

```
QY 661 YDLALDSGRPHSAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
DB 559 YDLALDPGRSIRAFDETKNNTRRTQVFGMLQKQETLKLILPDCVDSVSPILRLNY 618
QY 721 SLVGTPLSAFNGLRPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 780
DB 619 TLVGEPLRSGNLRPVLAEDAQRFMTAMPFPEKNCNDNICDDLSITWSMGLDTLVVG 678
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPLDLSYRKVSTLQNRQSRWRL-ACESASSTE 839
DB 679 GPQDFNMSVTLRNDGDSYGTQVTVYPSGLSYRKDSASQNPULKPFVFKPAESSSSSE 738
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLILKANTVSENMPRNKTE 899
DB 739 GHGALKSTTWNINHPFIPANSEVTNITFDVDSHASFGNKKLLKATVASENNMSRTHKTK 798
QY 900 FOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVLVP 959
DB 799 FOLELPVKYAVYVMTSDRESSIRYLNFTASEMTSKVIHQYQVNNLQGRSLPVSFVFWIP 958
QY 960 VRLNQTIVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVYVNCISIAVCQRIQCDI 1019
DB 859 VQINNVTVDHPQVIFESQNLSSACHTEQKSPHSNFRDQLERTPVLNCSVAVCKRIQCDL 918
QY 1020 PFGIOEENATLKGNSLDWYIKTSHNHLILVSTABILFNDSVFTLLPQOGAFVRSQTE 1079
DB 919 PPSNTQEIENVTILKGNLSDFWYIKTSHGHLILVSTABILFNDSAFALLPQGESYVRSKTE 978
QY 1080 TKVEPEVNPPLIIVGSSVGGILLALITAAALYKLGPKRQYKQNMWSSGGPPGASPO 1137
DB 979 TKVEPEVNPVPLIIVGSSVGGILLALITAGLYKLGFFKRYQKQNMWNEAAPQDAPPQ 1036

RESULT 3
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.4%; Score 3786.5; DB 6; Length 920;
Best Local Similarity 78.4%; Pred. No. 3.6e-272;
Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

QY 118 POKFPEALRGCPQOESDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTLFSLMQYS 177
DB 1 POKFPEALRGCPQOESDIAFLIDGSGINRLDQRMKEFVSTMVGMQFQKSKTLFALMQYS 60
```

```
QY 178 EBFRIHTFKERFONPNPRSLIKPITOLLGRTHRTATGLRKVYRELPNITNGARKNAFKIL 237
DB 61 EDFYHTFTFNDPKRNPSPKLLVRPILQLLGRTHRTATGIRKVVRELFSKSGARENAKIL 120
QY 238 FLITDGEKFGDPLGVEDVIPLEDRGVIRYVLGFGDAFSEKSRQELNNTVASKPRPDHVF 297
DB 121 VVITDGEKFGDPLGVEDVIPLEDRGVIRYVLGFGDAFSEKSRQELNNTVASKPRPDHVF 180
QY 298 QANNEALKTQVQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYD 357
DB 181 QVNFNEAVKTIQNLQEKTF AIEGTQTGSSSTSECEMSQEGFSAAITSNGLLSTVGSYD 240
QY 358 WAGGVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVLCAPRYOHIGLVA 417
DB 241 WAGGAPLHPKDRVFIINTTRVDSMDNDAYLGYAVEVILNRQAQSLVLCAPRYOHTGLV 300
QY 418 MFRQNTGMWSEANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQOVSV 477
DB 301 MFKQNSGAWEKWADIKGSIQYFGASLCSVDVNRDGSDDLVLIGAPHYVEOTRGQOVSV 360
QY 478 CPLPRGQARWQCDVAVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEENRGAV 537
DB 361 CPLPQG-RAKWCQCVILCGEQGHPNSRFGAALTALGDVNGDKLTDVAIGAPGEEENRGAV 419
QY 538 YLPHGTSGSGISPSHSQRIAGSKLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLL 597
DB 420 YLPHGTSELGISPSHSQRIAGSKLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLL 479
QY 598 RSQPVLRVKAIMEFNPREVARNVFECDVGVKKEAGEVRLVHVKSTDRDLRREGQIOS 657
DB 480 RSQPVLRVBAWVFAPREVARNVYECRQAAKTQIAGEVQVCLQVRKSTMDRLREGDTS 539
QY 658 VVYDLDLDSGRPHSAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRL 717
DB 540 IITYDLALDPGRPHRAVFEETKNNTRQTVLGLSRKCEHLALWLPDCVEDSVTPVLR 599
QY 718 LNPSLVGTPLSAFNGLRPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDC 777
DB 600 LNPSLVGKPASSFNGLRPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDC 659
QY 778 VVGGRPEFNVTVVRNDEGDSYRTQVTFPPLDLSYRKVSTLQNRQSRWRLACESAS 837
DB 660 VVGGRPDLKTLTVRNQGEDSYRTQVTFPPLDLSYRKVSTLQNRQSRWRLACESAS 719
QY 838 TEVSALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLILKANTVSENMPRNK 897
DB 720 TEESTALKSTSCSINHPIFPDENSEVTNITFDVDPDAFLGYKLLKANTVSENMPSSNK 779
QY 898 TEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVL 957
DB 780 TEFQLELPVKYAVYVMTSLVSTKYFNFTASEKTRHVIHQYQVFNGLQGRKLPISWFW 839
QY 958 VVRLNQTIVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVYVNCISIAVCOR 1017
DB 840 VVRLNQTIVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVYVNCISIAVCOR 899
QY 1018 DIPFGIOEENATLKGNSLF 1038
DB 900 DIPSGIOEELKVTILKNSLF 920

RESULT 4
Q81VA6
ID Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RC ISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237; AAH38237.1;
SQ SEQUENCE 1169 AA; 128521 MW; A17B484FEFC79EB6 CRC64;

Query Match 58.8%; Score 3457; DB 4; Length 1169;
Best Local Similarity 60.7%; Pred. No. 1.7e-247;
Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAVNOGSLVQCDYSTGSCPEI 60
Db 20 FNLDTTELTAFRVDAGFGDSVVQVANSVWVVGAPQKITAANQTGGIYQCYSTGACEPI 79

Qy 61 RLQVPPEAVNMSLGLSLAATTSPPQLLACGPTVHTQCTSENTYVKGCLCFGLSGNLURQOPK 120
Db 80 GLQVPPEAVNMSLGLSLAATTSPPQLLACGPTVHHECGRNMYLTGCLFLLGPT--QLTQR 137

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQKKSKTLFSLMOYSEEP 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMNFVRAVISQFQPSQFSLMQFSNKF 197

Qy 181 RIHFTFEFQNNPNRSLIKPIITOLLGRTHATGLRKVVRLEFNTGARKNAFKILFL 240
Db 198 QTHFTFEFRSSNPLSLASVHQLOGFTYATATQNVVHRLFHASVAGARDAAKILIVI 257

Qy 241 TDGKFGDPLGYEDVPELDRGVIRYVVGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
Db 258 TDGKFGDPLGYEDVPELDRGVIRYVVGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 317

Qy 301 NPEALKTVQNLREKIFAIEGTQGTSSSPHEMSQEGSAITNSGPLLSTVGSYDWAG 360
Db 318 DFDALKIQONLQKREKIFAIEGTQGTSSSPHEMSQEGSAITNSGPLLSTVGSYDWAG 377

Qy 361 GVFLTSKEKSTFNTMTVDSDMDAYLGVAATILNRVOSLVGLAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQVMDRSLYCYSTELALWKGVSLLVGLAPRYOHTGKAVIFT 437

Qy 421 QNTGMWESNANVKGITQIAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQVVCPL 480
Db 438 QVSRQWRMKAETVGTQISYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQVVCPL 497

Qy 481 PRGARWOCDAVLYGEGQWPGFRGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVLYF 540
Db 498 PRGWR-RWMCDAVLYGEGQWPGFRGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVLYF 556

Qy 541 HCTSGSGISPSHSORIASKLSPLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSORIASKLSPLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616

Qy 601 PVLRYKAIMFNPREVARNVFCNDQVVKGEAGEVVRVCLHVOKSTRDRLREGQTSVYT 660
Db 617 PVLWVGVMQFIPAEIPRSAFECREQVVSQTLVQSNICLYIDKESKNLLGSRDLQSSVT 676

Qy 661 YDLALDSGRPHRAVFNETHQSTRQTVGLGTQTCETLKLQPLNCIEDPVSPIVLRNLP 720
Db 677 LDALDPGRLSPRATFQETKRNLSRVRVLGLKAHCENFNILLPSCVEDSVTPITRLNF 736

Qy 721 SLVGTPLSAGNLRVLAEDAORLFTALFPPEKNCNDNI CODDLSTIFSMSLDCLVVG 780
Db 737 TLVUGPLLAFLNRLPMLAADAQRYFTASLPFFKNCADHI CODNLGISFSPFGLKSLVVG 796

Qy 781 GPREFNVTVVRNDEGDSYRTQVTFEFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 SNLELNAEVMWVNDGDSYGTITVTFSHPAGLSYRYVAEGQKQGLRSLHLTCDSPAVG-- 854

Qy 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 SQWTWSTSCRINHPIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSENNTPTRTSKTF 914

Qy 901 QLELPVKAVYVTVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
Db 915 QLELPVKAVYVTVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974

Qy 960 VRLNQVTVWDRPQVTFSENLSSTCTHYKERLPSHSDFLAELRAKAPVNCISIAVCORIQCDI 1019
Db 975 VELNQEAVWMDVEVSHPOPNPSLRCSSEKIPASPADFLAHIQKNPVLDCSIAGCLRPRCDV 1034

Qy 1020 PFGIOGEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQTE 1079
Db 1035 PFSVQOEELDTLKGNSLFGWVRQIILQKVSVVSAEITFDTSVYSQLPQOEAFMRAQTT 1094

Qy 1080 TKYEPFEPNPLPLIVGSSVGGLLLALITALYALYKLGFFKQYKDMWSE 1128
Db 1095 TVLEKYKHNPTELIVGSSIGLLLALITALYALYKVGFFKQYKDMWEE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
ID Q9QXH4;
AC Q9QXH4;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAF23492.1;
DR HSP; P11215; IBHQ.
DR MGD; MGI:96609; ItgaX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.1%; Score 3297.5; DB 11; Length 1169;
Best Local Similarity 56.2%; Pred. No. 1.2e-235;
Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAVNOGSLVQCDYSTGSCPEI 60
Db 20 FNLDAEKLTHFDHMGAEFGHSGVLYQDSSVWVVGAPKEIKATNQIGGLYKCGYHTGNCPEI 79

Qy 61 RLQVPPEAVNMSLGLSLAATTSPPQLLACGPTVHTQCTSENTYVKGCLCFGLSGNLURQOPK 120
Db 80 SLQVPPEAVNISLGLSLAATNPSWLLACGPTVHTTCRENIYLTGLCLFLLSSSPKQS-QN 138

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQKKSKTLFSLMOYSEEP 180
Db 139 FPTAQECPRQDQDIVFLIDGSGSISSTDFEKKLDFVKAVMSQLQRPSTFSLMQPSDYF 198

Qy 181 RIHFTFEFQNNPNRSLIKPIITOLLGRTHATGLRKVVRLEFNTGARKNAFKILFL 240
Db 199 RVHFTFNFNISTSPSLSGSVRLRGYTYTASAIKHVITELFTTQSGARQDQATKVLIVI 258

Qy 241 TDGKFGDPLGYEDVPELDRGVIRYVVGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
```


Db 259 TDGRKQDNISYSDVIMAEAAIIIRYAGVGFYNEHSGKELKAIASMPSEHYVPSVE 318
QY 301 NFALKTVQNLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSVDWAG 360
Db 319 NFDALKDIEQNLKEIFAIEGTETPSSSTFELEMSQEGFSAVTPDPGVLAGVGSFWSG 378
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHTGLVAMFR 420
Db 379 GAFLYPSNMRPTFINMSOENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 438
QY 421 QNTGMWESNANVKTGTCAGVFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGQGVSCPL 480
Db 439 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDSGLDLVLIGVPHYYEHTRGQGVSCPM 498
QY 481 PRGORARWOCDAVLGYGQGPWGRFGAALTVLGDVNGDKLTDAVAGPGEEDNRGAVYLF 540
Db 499 P-GVGRWHCGTTLHGEGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVYIF 557
QY 541 HGTSGGISPSHSQRIAGSKSLPRLOYFGOSLGGQDLTMDGLVDLTGAGQHVLRLRSQ 600
Db 558 HGASRQDIAPSPORISASQIPSRIOYFGOSLGGQDLTRDGLVDLAVGSKGRVLLRTR 617
QY 601 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRREGIOISVVT 660
Db 618 PILRVSPVTHFTPAEISRSYFECQEQVAPRQTLSDATVCLIHIESPKTQL--GDLRSTVT 675
QY 661 YDLALDSGRPHSAVFNENKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 FDALDHGRUSTRRAIFKETKTRALTRVKTGLNKHCSVKLLLPACVEDSVTPITLRLNF 735
QY 721 SLVGTPLSARGLRPLVLAEDAQRALTALPFFKXNCGNDNICQDLSITFSFMSLDCLVWG 780
Db 736 SLVGVPISSLQNLQPLAVDDQYFTASLPFFKXNCGADHLCQDLSLVFVGFPLDKTLVG 795
QY 781 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSVRKYSTLQ-----NORSQSWR 829
Db 796 SDLELVNVTVSDNGEDSYCTVTLFYPVGLSFRRAEQGVFLRKXEDQOMQRGQSHLH 855
QY 830 LACESASSTEVGALKSTSCSINHPIFPENSEVTENITFDVDSKASGNKLLKANVTSE 889
Db 856 LMCDD--STPDRSGLSTSCSRHVIFRGSGQMTFLTFDVSFPAELGRLLRLARVGE 913
QY 890 NMPRTNKTFFOLELPKVAVMVMTSHGVSTKYLNTASE--NTSRVMOHQYOVSNLQOR 948
Db 914 NNVPKPTTFFOLELPKVAVMVMTSHHDQFTKYLNFSEKSEKTSVVEHFRVNNLQOR 973
QY 949 SLPLSLVPLVRLNQTVMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCS 1008
Db 974 DVPVSINFVPIELKGEAVW-TVMVSHPNPLTQCYRNRLLKPTQFDLLTHMQKSPVLDCS 1032
QY 1009 IAVCQRIQCDIPFGIOEENATLKNLSFDWIKTSHNHLILVSTAEILFNDVSFTLLP 1068
Db 1033 IADCLHLRCDIPSLGIDELLYFLTKNLSFGMTSQTLQKVKLLSBAEITFNTSVYSQLP 1092
QY 1069 GQAFVRSOTETKVEPPEVNPPLVLVSSVGLLALLALITAAALKYKGFKKYKQWMSSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKVNPNVPLVGVSSVGLLALLAILTALYKAGFFKRYKEMLEE 1152

RESULT 6
Q9QE7
ID Q9QE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QE7;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietech G.,
RA Gallatin W.M.,
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;
Query Match 54.9%; Score 3225.5; DB 11; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.6e-230;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;
QY 2 NLDNTAMTFQENARGFGOSVVQLOGSRVVVVGAPQEIIVAAHQRSILYQCDYTGSCCEPIR 61
Db 21 NLDVEEPIVFRDEDAASFGQTVVQFGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCOPIV 80
QY 62 LQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVVKGLCFLFGSNLRQOPQKF 121
Db 81 LRSPLAEVNMISLGLSLVTATNNAQLLACGPTAQACVKNMYAKGSCLLLGSSL-QPIQAV 139
QY 122 PVALRGCPQEDSDIAFLIDGSGSIIPHFRNKELVSTIMEQKSKTILFSLMQVSEBFR 181
Db 140 PASMEPCPRQEMDIAFLIDGSGSINQRDPAQKDFVKALMGFASSTLTFLSLMQYSNLIK 199
QY 182 IHFTKEFONNPNRSLIKPITQLGRTHATGLKRVVRELNITNGARKNAFKILFLIT 241
Db 200 THFTTEFNILDDPSLDVPIVQLQGLTYTATGIRTVMEELPHSKNGSKRSKAKILLVIT 259
QY 242 DGEKEGDPGLGVEDVIPELDREGVIRYVLGFGDAFRSEKSRQELNIVASKPPRDHVFQANN 301
Db 260 DQKYRDPLEYSDVIPADKAGIIRYAGVGFDAFOETALKELNTIGSAPPQDHVKVCN 319
QY 302 FEAATVQNQLREKIFAIEGTQTSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAAG 361
Db 320 FAALRSIQRLQEKIFAIEGTQTSRSSSFQHEMSQEGFSSALTSQGPLGAVGVSFWSG 379
QY 362 VFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHTGLVAMFR 421
Db 380 AFLYPNTNRTPTFINMSQENVDNRDYLGYSTAVATFWKGVHSLILGAPRHQHTGKVIFTQ 439
QY 422 NTGMWESNANVKTGTCAGVFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGQGVSCPL 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSGLDLVLIGVPHYYEHTRGQGVSPVP 499
QY 482 RQARARWOCDAVLGYGQGPWGRFGAALTVLGDVNGDKLTDAVAGPGEEDNRGAVYLF 541
Db 500 -GVGRWQCEATLHGEGHPWGRFGAALTVLGDVNGDNLADVAIGAPGEESRGAVYIF 558
QY 542 GTSGSGISPSHSQRIAGSKSLPRLOYFGOSLGGQDLTMDGLVDLTGAGQHVLRLRSQ 601
Db 559 GASRLIEMPSQSRVTSQSLRLQYFGOSLGGQDLTQDGLVDLAVGAGQHVLRLRSLP 618
QY 602 VLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRREGIOISVVT 661
Db 619 LLKVELSTRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKGSFDLL--GVQGSVRY 676
QY 662 DIALDSGRPHSAVFNENKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSIVLRLNFS 721
Db 677 DIALDPGRILSRAIFDETKNCTLTGRKTLGLGDHCHETVKLLPDCVEDAVSPILRLNFS 736

QY 722 LVGTPLSAFNGLRPLVLAEDAORLFTALPFPFEKNCNGNDICODDISITFSPMSLDCLVVGG 781
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
737 LVRODASP-RNLHPVLAVGSQDHIITASLPFEKNCCKQELLCEGLGISFNFSGLQVLVVG 795
QY 782 PRFNVTVVRNDGEDSVRTQVTFPPFDLSYRKVSTIQNRSQSRWRLACESASSTVS 841
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
796 SPFLTIVTVWNEGSDSGTLVKFYYPAGLSYRRVTGTQ-QPHQVPLRLACEAPEAAQED 854
QY 842 GALKSTCSINHPIPEPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEFO 901
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
855 --LRSSCSINHPIPREGAKTTFMITFDVSYKAFGLDRLLURAKASSNNKPDNTKTAFO 912
QY 902 LELPKYAVYVMVTVSHGVSTKYNFTASENTSR-VMQHQYQVNSLQGRSLPISLVFLVPV 960
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
913 LELPKYTVYTLISQEDSTNHNFSSSHGRRQEAHRYRVNNLSPLKLAVRVNFVVPV 972
QY 961 RINQTVIWRDPOVTFSENLS--TCHTKERLPSHSDFLAELKAPVNVGSIACVQRIQCDI 1019
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
973 LLNGVAVMD--VTLSSPAQGVSCVSMKPPONPDFTQIORSSVLDCSIADCLHFRCDI 1029
QY 1020 PFFGQEBFNATLKNLSDFWYIKTSHNHLITVSTAEILENDSVFTLLPGOGAFVRSOTE 1079
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1030 PSLDLODELDFLRNLGFGWVSQTLQEKVLLVSEAITEDFTSVYSQLPQGEAFURAQVE 1089
QY 1080 TKVEPEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDM 1126
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1090 TTLEEVVVEPIFLVAGSSVGGLLLLALITVLYKLGFFKQYKEM 1136

RESULT 7

Q9WTV4
ID Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma B.Z., Teuscher C.;
RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1;
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWPA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 25.7%; Score 1512.5; DB 11; Length 1161;
Best Local Similarity 33.7%; Pred. No. 5.4e-103;
Matches 394; Conservative 219; Mismatches 464; Indels 91; Gaps 37;
QY 1 FNLDTENAMTFQENA-RFGQSVQVLOGSRVVVVGAPQETVAANQGSVQCDYSTGSCPE 59
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
24 YNLDTRPTQSFQAAGRIFGYOVQIEDG-VVVGAPGE---GDNTGGLYHCKTTSERCP 79
QY 60 IRLQVPEAVNMNLSGLSLAATTPPQLLACGPTVHTQCTSENTYVKGCLCLFSGNLRQOPQ 119

Db 80 VSLH-GSNHRSKVLGTMTLATDAAGSLSLACDPGLSRTCDQNTYLSGLCYLPQSGLEPML 138
QY 120 KFPALRGCEQSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTKLPSLMQYSEE 179
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
139 QNRPAYQECMKGVLDLFLFDGSDLRKOFKILEFEMKDMRKLSKLSNTSQFAAVQST 198
QY 180 FRIHFTKEF-QNNPNPRSLIKPITQLLGRTHTATGLRKVVVRELNFNTNGARKNAKILF 238
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
199 CRTEFTFLDYVKQKNKDFVLLGSVQPMFLLTNTFRAINYVAVHVKESGARPDATKVLV 258
QY 239 LLTDGEKF--GDPLGYEDVPELDREGVIRYVVGFGDAPRSEKSRQELNVTASKPPRDHV 296
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
259 IITDGEASDKGNISAAHD-----ITRYIIGIGHFVSVQOKTLHFIFASEPVEFV 309
QY 297 FOANPFALQVQNLREKIFATEGTQTGSSSSPEHEMSQEGFSAATITSGNPLLSVTGSY 356
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
310 KILDTFELKDLFTDLQRRYIATEGNRQDLTSFNELSSSGISADLSKHAVVGAGAK 369
QY 357 DWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGAAAA-IILNRVQSLVGLGAPRYQHIG 414
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
370 DWAGGLDLREDLQATFVGQEPDTSVDRGGLGYTVAMWTSRSTRPPLAAGAPRYQHV 429
QY 415 LVAMFR--QNTGMWESNANVKQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRG 472
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
430 QVLLFQAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDDGAEALLIGAPLFFGQRG 489
QY 473 GOVSVCPLPRQARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTVAICAPGEED 532
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
490 GRVETY--QRRQSLFEMVSELQDGPYPLGRFGAAITALTIDINGDRLTDVAVGAPLEE- 545
QY 533 NRGAVYLFHGTSGSGISPSHSHQSIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQOG 592
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
546 -QGAIVYFNGKPG-GLSPQPSQRIQGAQVFPGRMFGRSIHGVKDLGGDLADVVVVGEG 603
QY 593 HVLLRSQPLRVKAIMFNPREVARNVFECNDQVVGKQAG-BVRVCLHVQKSTRDLRL 651
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
604 RVVYLSRPVVDVVTLSFSPPEIIPVHEVECSAREEQKHVKLAKAFRIKPLTPQ--F 661
QY 652 EGQIQSVVTVDLALDGRPHSRVAFNETKSTRQTVGLTQTCETLKLQLPNCIEDPV 711
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
662 QGRULANLSYTLQDGHMRSGRGLPPGSHLSGNTSITP-DKSLDFHFFPFIQDLI 720
QY 712 SPIVLRNLNLSLV---GTPLSAFGN-LRPVLAEDAQRFTALPFPFEKNCNGNDICODDISI 767
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
721 SPINVSINFLSLEEGTFRQKGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCEANLTL 779
QY 768 TFSFMSLDCLVVGCP-----REFNVTVVRNDGEDSVRTQVTFPPFDLSYRKVSTLQ 821
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
780 SSPARS-----QPLRLMSASLAVENTLSNGEDAYVWRLDLDPRGLSFRKVEMLQ- 831
QY 822 QRSQSRWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTENITFDVDSKASLGNKL 880
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
832 --PHSRMPVSCSEL--TEGSLTLTKLVKNVSSSIFKAGQEVSVQVWMENTLNSWEDPV 887
QY 881 LLKANVTSEN-NMPTNKTPEQLPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMOHQ 939
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
888 ELNGTVHCENENSSLOEDNSAATHIPLYVNVILTKQENSTLYISFTPKPKTQVQHV 947
QY 940 YQV---SNLQORSPLISLVFLVPVRLNQTVIWRDQ--VTFSENLS-----TCHTKE-RL 989
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
948 YQVRIQPSAYDHNMP-TLEALVGP-----WPHSEDPITYTWSVQTDPLVTCHSEDLKR 1000
QY 990 PSHSDFLAELRKAPVNCSTIAVCORIOCDIPFFGQEBFNATLGNLSFDWYIKTSHNHL 1049
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1001 PSSE---AEQCLPGV-----QFRCPVIF---RREILLQVGTGTVSELSKEIKAS-STL 1045
QY 1050 LIVSTAEILFNDSVFTLLPGQAFVRSQETETKVPFPEVNPPLIVGSSVGGLLLLALIT 1109
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1046 SLCSLSVSFNSSKHFLYGSKA-SEAQVLVKVDLIHEKEMLVHVVLVSGIGGLVLLFLIF 1104
QY 1110 AALYKLGFFKQYKDMN-SEGPPGAP 1136
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

Db 118 SYKTEFDSYVVRKDPDALLKHVHMLLLTNTFGAINVYVTEVFREELGARPDPATKULI 177
Qy 239 LITDGEKFGDPLGYEDVIFELDREGVIRYVYVFGDAPRSEKSRQELNTVASKPPRDRHVFQ 298
Db 178 IITDGE--ATDSNIDAAND-----IIRYIIGIKHFQTKESQETLHKFASKPASEFVKI 230
Qy 299 ANNFEALKTVQNLREKIFAIEGTQTGSSSPHEMSQBFSAITNSGPLLSTVGSYDW 358
Db 231 LDTFEKLKOLFTLQKIVIEGTQKDLTSFNMELSSGISADLSRGHAVVAVGAKOW 290
Qy 359 AGGVF-LYTSKESKSTINTRVDSMDNDAYLVGAAA-IILNRVQSLVAGAPYOHIGLV 416
Db 291 AGFLODKADLQDDTFIGHEPLTPVRAGVLTVTWLPQRKTSLLASGAPRYQMGV 350
Qy 417 AMFR--QNTGMWESNANVKTGIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQRGGQ 474
Db 351 LLFQEPQGGHWSQVTHGTQIGSYFGGELCGVDVDQGETELLIGAPLFYGEQRGR 410
Qy 475 VSCPLPRGRARWQCDV--LYGEOGPWGRGAALTVLGVNGDKLTDVAIGAPGED 532
Db 411 VFYI-----QRQLGFEEYSELQDQFGYPLGRFGEAITALTDINGDGLVDVAVGAPLEE- 464
Qy 533 NRCAYLHGTSGSGISPHSQRISAGSKLSPRQYFGOSLSSGQDILTMGLVDLTVGAOS 592
Db 465 -QCAVYFNGRHG-GUSPQSQRIEGTQVLSGIOWFGRSIHGVDKLEGDLADVAVGES 522
Qy 593 HVLLRSQPLVRKALMEFNPREVARNFECNDQV--KGKEAGEVRCVCLHVKQSTRDLR 651
Db 523 QMIVLSRPRVDMVTLMSFPAIPVHEVECSYSTSKMKEGYVITICQI-KSLIPQF- 580
Qy 652 EGOIQSVVYVDLALDSGRHSAVFNKTNSTRQTOVLGLTQCTETLKLQIPNCIEDPV 711
Db 581 QGRVANLTYTLQDGHRTRRRGLFPGRRHRLRNIAVT-TSMSCDTSFSPHPCVQDLI 639
Qy 712 SPVLVLNLSL-----VCTPLSAPGN-----LRVLAEADAQRLFTALFPREKACGNDNIQ 762
Db 640 SPINVSLSLWEEBETPDQRAGKDIPIRLPSLHSETWEI-----PFEKNGEDKKCE 694
Qy 763 DLSITFSFMSLCLVGVGPREFNVTVRNDCEDSYRTQVTPFFPLDLISYRKVSTLQNG 822
Db 695 ANLRVSFSPARSALALTAFASSLSVELSLNLEEDAYVQLDLHFPGLSFRKVEML--- 751
Qy 823 RSORSRLACES--ASSTVSGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKL 880
Db 752 KPHSQIPVSCPEELPEESRLLSRAL--SCNVSSPIFKAGHSVALQMMFNLTVNSSWGDSV 808
Qy 881 LKANYTSENN---MPTNKTEFQLELPVKYAVVWVTVSHGVSTKYLNFTASENTRVM 936
Db 809 ELHANYTCNNEDSDLLIEDNSATTI---IPILYPINILIQOEDSTLYVSFTPKGPKHQV 865
Qy 937 QHOYQV---SNLQORSIP-LSLVFLVPLVRLNQVWDRPOVTFSENLSSTCHTK--ERLP 990
Db 866 KHWYQVRIQPSIHDHNIPTLEAVVGVPOPPSEGPITHQWSVQMEPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVNVCSIAVCORIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNH 1048
Db 924 DAAEPCLPGALFPCPVV-----FRQELVQVIGTLELVGIEAS-SM 964
Qy 1049 LLIVSTAELFNDSVFTLLPGQAFVRSQTETKVEFEVNPPLPLVGVSGVGLLLALI 1108
Db 965 FSLCSSLISFNSKHFHLYGSNASL-AQVVMKVDDVYVYQKMLYLVLSGIGGLLLLI 1023
Qy 1109 TAALYKLGFPFKROYKDMSEG-CPPGAEP 1136
Db 1024 FIVLYKVGFPKRNKEMAGRGVNPPIP 1052

RESULT 11

Q98TF0

ID Q98TF0 PRELIMINARY; PRT; 1187 AA.

AC Q98TF0;

DT 01-JUN-2001 (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID:7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1; -
DR HSSP; P20701; 1LFA.
DR InterPro; IPR001969; Asprotease site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match

22.6%; Score 1329.5; DB 13; Length 1187;

Best Local Similarity 31.0%; Pred. No. 2.2e-89;

Matches 369; Conservative 210; Mismatches 456; Indels 157; Gaps 43;

Qy 1 FNLDTENAMTFOENARG-FGOSVVQLO-GSR--VVVGAQOEIVAAQORSLYQCDYSTGS 56
Db 32 FNIDTEHPURNGTDEDFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSCTADLOS 89
Qy 57 CEPRLQVP-----VEAVNMSLGLSLAATTPSPQLLACGPTVHTQTSNTYVYVGLCLFLGS 112
Db 89 CK--RLQRPGSSEVFFGMSAAVSSAULTS-----CSPYFAHECDGNSYLVGVCYQFNS 140
Qy 113 NLRQOPQRPPEALRQCPQEDSDIAFLIDGSGSIIIPDPRMKELVSTIMEOLKSKTLFS 172
Db 141 SL-QAVSNFTAAAYQCSKREVNLFVLFDFGSSMKAVEFDMKNFIDVMKLSNSSIKFA 199
Qy 173 LMOYSEERIHTTFKEFQNNPNRSLI KPIITOLLGRTHATGLRKVVRELFN-ITNGARK 231
Db 200 AVQFSTEIRTVDFDNDYQNGSAEELMKE-RHMKSLTNTYKAINYYLVKNVLSVSSGADP 258
Qy 232 NAFKILFLITDCEKFGDPLGYED--VIPELDREGVIRYVVLGFGDAPRSEKSRQELNTVAS 289
Db 259 NAKALVIITD---GDSDDNDYLNLCDEQNILRYIIVG-----KVDLTTLTQLAA 309
Qy 290 KPPRPHVFOANNFEALKTVQNLREKIFAIEGTQTGSSSPHEMSQBFSAITNSGPL 349
Db 310 EPKLANTFYIOBYNGLKGLDLNLQKKIYNIEGSKAHRDQKELSSQSGFVVYQESVI 369
Qy 350 LSTVGSYDWAGVFLYT---SKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVLG 406
Db 370 VGSVGSNDNRGALYEVMSGSKFRQTEITDPVAVN---KDSYMGYSTVLGMRRHGVSLFFSG 426
Qy 407 APYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHY 466
Db 427 APRAEHTGLVLTFTKNESTWTVMRNINGEIQISYFGASLSDVDSDGSDFLVLVGAFLP 486
Qy 467 YE-QTRG-QQSVCPPLRQQRARWQCDVLYGEQ--GQWGRFGAALTVLGVNGDKLTD 522
Db 487 YSQPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDINGDGLSD 535
Qy 523 VAIGAPGEDNRGAVLPHGTSGSGISPSHS--ORTAGSKLSRPLQFGOSLGGQDLTWD 581
Db 536 VAVGAPLE--NEGVVYIYLGDTGHGINPEHAPQIPARSVLFGLOQFGVSLSGQMDMND 593

Qy	582	GLVDLTGAQGHVLLRLRSQPVLVKTAIMEFNPREVARNVPECNDQVVKGEAVRVCLH	641
Db	594	NLPDVIIVTGGIVLLNARVMSVSAQLSPNPWEISLNYECPGS--NAPNAPNLTSCT	651
Qy	642	VQKSTRDLRREGQIQSV--VTYDLALDSGRPHRASVPNETYNSTR--RQTQVGLTQTCE	697
Db	652	VTERTSS--TGSLEKKLVNLSLNLNDVVRGMRGFFDPMDDSSRLLQOSVLLDSSCS	708
Qy	698	TLXLQLPNCIEDPVSPIVLRNFS---LVGTPLSAFGLNRPVLADQAQLFTALPFPEK	753
Db	709	NFSIFMLRCVADTVSPUKIRMNFSOTOMLSGNSLAVL-DIQSSTEEVEVL-----	761
Qy	754	NCNGNDNICODDLSTTFSEMSLDCVLVGGPREFNVTVVRNDGDSYRTQVTFPPFLDLSY	813
Db	762	NC-NSNSCVADLKLNFSEFTN-DTLVVENQAHTVLYSLANPGDSDYNTSVLHYPEGLSL	819
Qy	814	RKVSTLQNRQSRWRLACSSASTEVSGALKTSCSINHPIIPENSEVTFNITFDV---	870
Db	820	SKFDAIKPSRTR-----SSCDRDSGATNRTTCSINLPVRSRGTTTQFLGTFRVTKW	871
Qy	871	DSKASLGNKLLKANVTSENNMPRTKTEFQLELPVKYAVMVVTSHG--STKLYNFTAS	929
Db	872	DYDWSDRMWTITANSNNGNM---SMSVRRSIPVQFAVELAISLVAEDSVTYLNFSE	928
Qy	930	ENTSRVMOHOVQSVNLCGSLPISLFLVPVRLNQTVIWRPQVTFSENISSTCHTKERL	989
Db	929	DRGPKPLNIYKVNGLGKLPVSVTLSPCQ-----TTHVTLTPHNFMSQ	974
Qy	990	PSHSDFLAELRKAPVVC-----SIAVCQRIOCDIPFFGQEE	1027
Db	975	EVVHSFISSYHQ--IIMCLLNKHLFFSPELSAVQTRTTGRSLWC-----VSSISTGEI	1026
Qy	1028	FNATLKGNL-----SFDWYIKTSHNHLIIVSTAELFNDSVFTLLPG--QGA	1072
Db	1027	FRSSV--NLMAEAVLQNVKEYESKYSFY-EPRDRDVFNIS-AELNFTNSRYNOSTGLKYN	1082
Qy	1073	FVRSQTEKVEPPEVPNPLPIVGVSSVGGLLLALITAAALYKLGFFKRYQKD	1124
Db	1083	PHRSQTEVKVEFVPPSRMLIVCTGAVGGFFFLIILILLKCGFFKRNRPD	1134
RESULT	12		
Q8HZV0			
AC	Q8HZV0	PRELIMINARY;	PRT; 927 AA.
DT	01-MAR-2003	(TReMBLrel. 23)	Created
DT	01-MAR-2003	(TReMBLrel. 23)	Last sequence update
DT	01-MAR-2003	(TReMBLrel. 23)	Last annotation update
DE			Lymphocyte function-associated antigen 1 (Fragment).
OS			Bos taurus (Bovine).
OC			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC			Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC			Bovidae; Bovinae; Bos.
OX			NCBI_TaxID=9913;
RN			[1].
RA			SEQUENCE FROM N.A.
RA			Thumbkat P., Kannan M.S., Maheswaran S.K.;
RT			"Sequence of the alpha subunit of bovine lymphocyte function-
RT			associated antigen 1.";
RL			Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF440778; AAN63636.1;		--
FT	NON_TER	1	1
FT	NON_TER	927	927
SO	SEQUENCE	927 AA; 102523 MW;	02E2CF09917081EC CRC64;

Qy	234	FKILFLLTGBEKGDPGLGVEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELINTVASKPRR	293
Db	64	TKVLIITIDGK---PPTNTLLMRPKTSS-----RSLLGIGNFKTKESQBALHQFASKPVE	116
Qy	294	DHVFOANNREALXTQONLREKIFAIEGTQTGSSSFEHEMSQEGFSAITNSGNPLLSV	353
Db	117	EFVKILDTFEKLDKLFTELOKKIYVIEGTSKQDLTSFNMELESSGISADLSEGHGVGVAG	176
Qy	354	GSYDWAAGGVF-LYTSKEKSTFFINMTRVDSMDMDAYLGAAAA-IILNRVOSVLVLGAPRYQ	411
Db	177	GAKDWAGGFLDKADLKSSTFVUGNEPLTVESRAGYLGYYTVTRPSRGTMSLLATGAPKYQ	236
Qy	412	HIGLVAMFRQ--NTGHWESNANVKTQIGAYFGCASLCSDVDVDSNGSTDLVLICAPHYEQ	469
Db	237	HVGRVLLFOQPRKGPMWSQIQEIDGQIGSYFGBELCGVDVORDGETELLTAAPLYYGE	296
Qy	470	TRGQVSVCPPLRGQARWOCDAVLXGEOQPMGRFGAALTVLGDVNGDKLTDAVLCAPG	529
Db	297	QRGRVPIY---QKIQLEFQMVSELOGETGYPLGRFGAAIAALTIDINGDELTDVAVGAPL	353
Qy	530	EEDNRGAVLYFHTGSSGTSPSHSQRIAGSKLSPRLOYFQGSLSGGQDLTMDGLVDLTVG	589
Db	354	EE--QGAVYIFNGQG--GLSPRSQRIEQTQMFSGIQWFGRSIHGVKDLGGDGLADVAVG	410
Qy	590	AQGHVLLRSQPVLRVKAIMEFNPREVARNVFCNDQVVKGEAG-EVRVCLHVQK--ST	646
Db	411	ABGQVILVSRPVDIITSVSFPAEIPVHEVECSYSTSNQKKEGVNLTVCFOVRSLSLT	470
Qy	647	RDRUREQIQSVVITYDLALDSGRPHSRAVENETKNSTRQTQVLGLTQTCTETLKLQLPNC	706
Db	471	----FOGHLVANLTYTLQDGHRTSRGLFPGGKHKHLGNTAVTPV-KSCFVPMFHPIC	525
Qy	707	IEDPVSPIVLRNPSL---VGTPLS--AFGNLRPVLAEADAQRULTALFPEKKNCGNDNIC	761
Db	526	IQDLISPINVSLSYSLMEBSGTPRDPRALDRDIPPLIKPSPHLETKEIPFEKKNCGEDKNC	585
Qy	762	QDDLSTIFSMSLDCLVVGGRPRENVTVTVRNDGEDSYRTQVTFPPRLDLSYRKVSTLQW	821
Db	586	EADLKLAFSDMRSKILRLTSPASLSUSVRLTURTAEADYVWQVTLSPFQGLSFRKVEL--	643
Qy	822	QRSQSRWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLL	881
Db	644	-KPHSHVPVGCCELPEEAUVHS-RALSCNVSSPIFGEDSMVDIQVMPNTLQKGSWGFIE	701
Qy	882	LKANVTS-----ENNMPRTNKTETQELPVPKYAYVMVTSHGVSVKYLNFTASENTS	933
Db	702	LQANVSCNNEDSSILLEDNATS-----IPVMYPINVLTKQENSTLYISTPKSPRI	754
Qy	934	RVMOHQYOV----SNLQSRSLPISLFLVPVRLNQTVI---WD---RPQVTFES-ENLSST	982
Db	755	HHVHHIYQVRIOPSNYDNMP-PLBALVVRVPRVHSEGLTHKWSIQMPEPPVNCSPRNLESP	813
Qy	983	CHTKERLPSSHSDFLAEALRKAPVNVNCISIAVCQRIQCDIPFGIOEBFNATLKNLSFDWYI	1042
Db	814	SDEAE-----SCSFCT--EPRCPIDF---RQEILVQVNGMVELRGTI	850
Qy	1043	KTSHNHLLIVTASILFNDSVFTLLPGOGAFVRSQETETKVEPPEVPNPPLPIVCSSVGG	1102
Db	851	KAS-SMLSLCSSLAISFNSSKHFLHGRNASM-AQVVMKVLDLYVEKMLYLYLVLSGGGL	908
Qy	1103	LLLLLITAALYKLGFFKR	1120
Db	909	LLLLLIFIALYKVGFFKR	926
RESULT 13			
ID	O88340	PRELIMINARY;	PRT; 1167 AA..
AC	O88340;		
DT	01-NOV-1998	(TReMBLrel. 08, Created)	
DT	01-NOV-1998	(TReMBLrel. 08, Last sequence update)	
DT	01-MAR-2003	(TReMBLrel. 23, Last annotation update)	


```

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02234; VWFA; 1.
DR SS SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.4%; Score 1142.5; DB 11; Length 1167;
Best Local Similarity 28.7%; Pred. No. 1.7e-75;
Matches 354; Conservative 213; Mismatches 461; Indels 205; Gaps 43;

QY 1 FNLDENA--MTQENARGFGSVQVQLQGRVVGAPQEIIVANORGS-----LYQCDY 52
DB 20 FNDVDWAWYALQCPAPVLSLLHLDPSN-----NQICLLVARSSNRNTAALYRCAL 74
QY 53 STGSCPIRLQVPEAVNNSLGLSLAATT--SPQLLAC-GPTVHQTCSNTVYKGLCLFL 109
DB 75 SI-SPDEIACQ-PVEHICNPKRGYQGVTLVGNHNGVLVCIQVQARFRSLNSELTCACSL 132
QY 110 FGSNLRCQPKPFPEARLG-----C-----PQE 131
DB 133 LTPNLDLQAAQYFSDLEGFLDPGAHVDSGDYCRSKGSGTEEEKSARRRTVEEEDDED 192
QY 132 DSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQL--KSKTLFSLMOYSEEFRIHFTKEF 189
DB 193 GTEIAVLDSGSGIERSDQAKNFISTWMRNEYKCFECNFALVQYGAIVQTEFDLQES 252
QY 190 QNNPNRSLIKPTQLLGRHTATGLRKVVRELFNITNGARKNAFKILFLLTDTGEKFGDP 249
DB 253 RDINASLAKVQSVQVKEVTKTASAMQHVLDNIFIPSRGSRKALKVMVVLTDGDI FGDP 312
QY 250 LGVEDVIPELDRGVIRYVLGFGDAPRSEKSRQELNTVASKPPRDHVFQANNPEALKTVQ 309
DB 313 LNLTTVINSFKMGVVVRFAIGVDGAFKNNTVRELKLIASDPKAEHTFKVTNYNSALDGLL 372
QY 310 NOLREKIFAIEGTQOTGSSSEFHEMSQEGFSAITSNGP--LLSTVGSYDWAGGVFLY-TS 367
DB 373 SKLQQRIVHMEGT---VGDALQYLAQTGFSAQLDKGQVLLGTGAFNWSGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGYAAAILNRVQSLVGLAPRYQHIGLVAMFRQNTGM 425
DB 430 NGRGCFNLQNTAKEDSRVTQVSYLGYSLAVLHKAHGISYVAGAPRHKLKRGAVPELRKEDR- 488
QY 426 WESNA---NVKGTQICAYRGASLCSDVDVDSNGSTDVLVIGAPHYHETQGGQVSVCLPR 482
DB 489 -EDAFVRRIEGQMGSYFGSVLCVPDIDMDGTDFLLAAAPYHIRGEEGRVYVQVPE 547
QY 483 GQARWQCDVAVLGEQGMGRFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSCHPLGTLNSRCGFMAAVGDIINQDKFTDVAIGAPLEGFAGDGAASYGS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGQDLTMDGLVDLTGVAQGHVLL 596

```

```

DB 607 VVIYNGHSG-GLYDPSQOIRASSVAGSLHYFGMSVSGGLDFNGGLADITVGSRDSAVV 665
QY 597 LRSQPLRVKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LBSRPVDLTVSMTEFP-----DALPMVFIGKM--DVNLCFEVDSSVASEPGLREM 715
QY 654 QIQSVVTVDLALDLSGRPHSRVFNETKNSTRQTQVLGLTQC----- 696
DB 716 FLNFTVDV-----TKQRQLQCEDSSGCCSCLRKMKNGSGSLCBHFPLI 760
QY 697 ETLKLQLPNCIEDPSPVILRLNFSLVGTPLPSAFGNLR-----PVLAEQAORLFTALF--P 750
DB 761 STEEL-----CEDCFSNITIKVYE-----FOTSGRRDYNPTL--DHYEPSAIFOLP 809
QY 751 FEKNGCNDNICODDLSITFSFMSLDCLVGGVPREFNVTVTVRNDGEDSVRTQVTFEFPDLD 810
DB 810 YEKCKKNVFCIAEIQLTIN-ISQELVVGVTKEVTMNLISLTSNGEDSYMTNMLNYPN 868
QY 811 LSYRKVSTLQNRQSRQSWRLACESASSTEVSGKSTSCSINHPIFPENSEVTFNITFDV 870
DB 869 LQFKKI-----QKPVSPDVQCDPKPV---ASVLWMCKIGHPIL-KRSSVNVSVTQQL 918
QY 871 DSKASLGNKLLKANVTSNNMPTNKTEFQLELPKYAVYVYVTSYHGVSTKYLNTASE 930
DB 919 EESVFPNRTADITVTISNSNEKSLARETR---SLQRFHAFIAVLSR--PSVYMN--TSQ 971
QY 931 NTSRVYMQHYQVSNLQGRSLPISLVFLVPLRNQTVIWDPROVTFSENLSST-----CHT 985
DB 972 SPSDHKEFFNVHGENLFGAVFQLOICVPIKLQDF-----QIVRVKMLTKTDHTECTQ 1025
QY 986 KERLPSSHDFLAELKAPVVMNCISIAVCORIQCDIPFFGIQEEFNATLKNLSFDWYIKTS 1045
DB 1026 SOEPACGSDPVQGVHKEWHSVVCAL-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILFNDVSFTLLPGQGAFFVRSQTETKVEFF-----EVPNPLPLIV 1095
DB 1066 HTKQLLRDVSELPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEEETRSLPLII 1119
QY 1096 GSSVGGLLLLALITAAALKYKGFQYKDMNSE 1128
DB 1120 GSSIGGLLVLVIIAILFKCGFFKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

```



```
DR SMART: SM00191; Int alpha; 4.
DR SMART: SM00327; WVA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON TER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;

Query Match 18.7%; Score 1097; DB 11; Length 1167;
Best Local Similarity 27.4%; Pred. No. 4.1e-72;
Matches 330; Conservative 205; Mismatches 427; Indels 244; Gaps 37;

QY 47 LYQCDYSTGSCPIRLQVPVEAVNNSLGLSLAATTSPQLLACQPTVH-----94
DB 67 LYQC-----ATSLIPDEIGCQPVHEHILMPKGRYQGV 98
QY 95 -----QTCSENTYVKGCLCFGLFOSNLRAQOQKPEALRG-----127
DB 99 LVRNHNGVLVCIQVOSRKRPRSLNSBELTGNLDLQAOAYFSDLEGVLDLGASVNS 158
QY 128 -----CPOED-----SDIAFLIDGSGSIIPHDPRMKELV 157
DB 159 GDYVSGKSGTGEETKSARRLRQRAVEDEEBAAGTEIAIYLDGSGSTEPSDFOKAKDFI 218
QY 158 STIMEQL---KSKTLFSLMOYSEEPRIHFTPEFQNNPNRSLIKPITQLLGRTHRTATGL 215
DB 219 STMRNFYKCFECNFALVOYGGVIQTDFDLDSRDINASLAKVQSIYOVKEVTKTASAM 278
QY 216 RKVRELNIYNGARKNAFKILFLTDGKFGDPLGYEDVLPEDREGVIRYVLFQGDAP 275
DB 279 QHVLNIFIPRSGSRKALKVMVLTDGDIIFRDLPLNLTVISSKMQGVVRFAGVGNAF 338
QY 276 RSEKSRQELINTVASKPPRDHFOANNFEALKTQVONLREKIFAIEGTGTGSSSFEHMS 335
DB 339 ENNTYRELKLIASDPKAHTFKVTNYSALDGLLSKLQRIIHMEGT---VGDTLQQLA 395
QY 336 QEGFSAITNSGP-LLSTVGSVDWAGGVFLY-TSKEKSTFINMTRVDSMDNDA---YLG 390
DB 396 QTGFSAQILDKGVLLGTVGAFNWSGGALLYNTQNGRFLNQT-AKEDFRAAQVSYLGY 454
QY 391 AAAIILNRVQSLVGLAPYQHIGLVAMFRONTGWESNANVKGTOICAYFGASLCSVDV 450
DB 455 SVAALHKAHGSYVAGAPRHKLRGAVFELOKEDGETPMRIEGEQMGSYFGSVLCPVDI 514
QY 451 DMSGSTDLVLIGAPHYEQTRCGQVSVCPPLPRGQARWQCDVLYGEOQPGWRFGAALT 510
DB 515 NMDGITDPLLVAAAPYHINGEGRVTVYTRV-HEQDAPFSLVTLTSGYFGLTSSRFGPAMA 573
QY 511 VLGDVNGDKLTDVAICAP-----GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR 564
DB 574 AVGDINQDKFTDVAIGAPLEGFGAGDGASVGSVYIYNGHSG-GLHASPSQIRASSVALG 632
QY 565 LOYFGSLSGGQDLTWDGLVLTGAQGHVLLRSQPVLRVKAINFEFNPREVARNVPECN 624
DB 633 LYVFGVSGGLDFSGDDLADITVGSQDVAVVLRSPVVDLTVSMTFTPDALP-----685
QY 625 DQVVKGEAGEVRVCLHWKS---TRDLREQIOISVVTYDLALDSGRPHSAFVNETKN 681
DB 686 ---MAFKQKMDVELCFKVDSSAVPSEPLRGMSLNFVTVDVDV-----TKQ 727
QY 682 STRRQ-----TQVLGLTQTCTETKLQLPNCIEDPVSPIVLRNLFS 721
DB 728 KORLOCADRSCCOSCLMKWSGSSSUCHEHGLLSTBEEL-----CEDDCFSNITIKVSVE 781
QY 722 LVGTPLSAFGNLRPLVAEDAQRFLTALF--PEKNGCNDNIQDDLSITFFSPMSLDCLVV 779
DB 782 F-QTSEERNHPNPII--DHYKEPSAIFQLPVEKDKCKNVFCIAETQLTALTSQOD-LVV 837
QY 780 GGPREFNVTVYRNDEGDSYRQVTFPFLDLISYKVKVSTLQNRORSWRHLACESASSTE 839
DB 838 GITKEVTWNISLNSGDSYMTNMAIYNPRNIQFKI-----QKPLSDIQDDPKPV- 890
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTE 899
```

```
891 --ASVLVNMCKIGHPIIL-KRSSVNVSVTWQLEESIPNRTADITVITISNSNEKSLARETH 947
900 FOLELEPKYAVVYVMTSHGVSTKYLNFNTASENTRVMOHQYQVSNLQORSLSPLSLVFLVP 959
948 ---SLOFRHAFIATLSR---PSVWYMN--TSQSSSDHKKEFFNVHGENHFGAVFQLQICVP 1000
960 VRLNQTVWDRPOVTFSENLSST-CHTKERLPSHSDFLAELRKAPVNVVCSIAVCQRIQCD 1018
1001 ITRDL-----QIIRVGHLTQTAHT-----ECTQSQSEPTCGSDPVQNV---1039
1019 IPPFGIQEERNATL-----KGNLSPDWIKTSHNHLLI--VSTAEIL-----PNDVSFTL 1066
1040 -----EMHVSIVCAITSNKENVTVAELSMGHTKQLLRDISLQILGEISFNKSLYEG 1092
1067 LPQGFVRSQETKVEPF-----EVPNPLPLIVSSGVGGLLLALITALYKLGFPFKROY 1122
1093 LNAE-----NHRKTIVIFLKEEKPHSLPLIIGSSIGGLLVLLVIAILFKCGPPKRY 1146
1123 KOMMSE 1128
1147 KQLNLE 1152

RESULT 15
O42094 PRELIMINARY; PRT; 1171 AA.
AC O42094;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE ALPHA1 integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
the alpha1 integrin gene";
RL J. Biol. Chem. 272:26643-26651(1997).
DR EMBL; AB000470; BAA23160.1; -
DR EMBL; AB000471; BAA23161.1; -
DR HSSP; P17301; IAOX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int. alpha; 4.
DR SMART; SM00327; WVA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Query Match 17.6%; Score 1034; DB 13; Length 1171;
Best Local Similarity 27.8%; Pred. No. 2e-67;
Matches 338; Conservative 213; Mismatches 513; Indels 152; Gaps 46;

QY 2 NLDTENAMTFOENARG-FGQSVVQL---QSGRVVVGAPQEIIVAAQNQSLVOCDSYTGSC 57
DB 19 NVDVKNAMTFSGLPDMFGYTVQQYENEBEGKWLIGSLVGVQPEKRTGDVTKCPVRDSQ 78
QY 58 EP-IRLQVP-----VEAV--NMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLC 107
DB 79 SPCKLNLPDATSVPNVMEVKENMTLGTTL-VTPNPKGGFLACGPLYAYKCGRLHYTTGVC 137
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKRQYKDMMSGGPPGAEPO 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	19	Human Beta-integrin
2	5839	99.3	1153	21	Human CD11b protei
3	5839	99.3	1153	23	Human Beta2 integr
4	5839	99.3	1153	23	Human integrin 1 a
5	5839	99.3	1153	23	Integrin Mac-1 alp
6	5829	99.1	1153	11	Alpha subunit of M
7	3437	58.5	1163	11	p150.95 alpha subu
8	3423	58.2	1163	19	Human Beta-integrin
9	3423	58.2	1163	21	Human CD11c protei

10	3423	58.2	1163	23	ABG61470	Human Beta2 integr
11	3421	58.2	1163	24	ABU07406	Protein differenti
12	3388	57.6	1161	16	AAW78166	Human Beta-2 integ
13	3388	57.6	1161	18	AAW23049	Human Beta 2 integ
14	3388	57.6	1161	19	AAW72825	Human alpha-d. Ho
15	3388	57.6	1161	19	AAW65089	Human Beta-integri
16	3388	57.6	1161	19	AAW57491	Human beta2 integr
17	3388	57.6	1161	20	AAW73342	Human alphad prote
18	3388	57.6	1161	21	AAW07359	Human alpha d clon
19	3388	57.6	1161	23	ABG61468	Human Beta2 integr
20	3372.5	57.4	1161	18	AAW23064	Human beta 2 integ
21	3372.5	57.4	1161	19	AAW72837	Human alpha-d deri
22	3372.5	57.4	1161	19	AAW65106	Human Beta-integri
23	3372.5	57.4	1161	20	AAW73343	Human alphad prote
24	3372.5	57.4	1161	21	AAW07376	Human alpha d prot
25	3372.5	57.4	1161	23	ABG61485	Human Beta2 integr
26	3215.5	54.7	1161	16	AAW78169	Rat alpha-d subuni
27	3213.5	54.7	1161	18	AAW23062	Rat beta 2 integri
28	3213.5	54.7	1161	19	AAW72824	Rat alpha-d #1. R
29	3213.5	54.7	1161	19	AAW60004	Rat alpha d polype
30	3213.5	54.7	1161	21	AAW07374	Rat alpha d protei
31	3213.5	54.7	1161	23	ABG61483	Rat Beta2 integrin
32	3206.5	54.5	1161	19	AAW65104	Rat beta-integrin
33	3206.5	54.5	1161	20	AAW73345	Rat alphad protein
34	3201	54.4	1161	18	AAW23061	Mouse beta 2 integ
35	3201	54.4	1161	19	AAW72836	Mouse alpha-d #2.
36	3201	54.4	1161	19	AAW65103	Mouse beta-integri
37	3201	54.4	1161	19	AAW60003	Mouse alpha d poly
38	3201	54.4	1161	20	AAW73347	Mouse alphad prote
39	3201	54.4	1161	21	AAW07373	Mouse alpha d prot
40	3201	54.4	1161	23	ABG61482	Mouse Beta2 integr
41	3197	54.4	1161	16	AAW78168	Mouse alpha-d subu
42	3192.5	54.3	1151	18	AAW23059	Rat beta 2 integri
43	3192.5	54.3	1151	19	AAW72834	Rat alpha-d #2. R
44	3192.5	54.3	1151	19	AAW65101	Rat beta-integrin
45	3192.5	54.3	1151	19	AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; Protein; 1153 AA.

XX AAW65090;

XX 28-SEP-1998 (first entry)

DE Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-0485618.

PR 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 21-DEC-1994; 94US-0286889.

XX (ICOS-) ICOS CORP.

XX Gallatin WM, Van DER VIEREN M;

PI

XX

CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11b. This
CC sequence was used in an alignment to identify a novel beta2 integrin
CC alpha subunit: alpha d (AAAG6014 and AAB07359). The present sequence has
CC approximately 60% identity to the protein sequence of alpha d. The
CC Alpha d gene and protein may be useful in therapy for diseases linked
CC to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis,
CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
CC antibodies may be used in the inhibition of macrophage infiltration at
CC the site of a central nervous system injury. The monoclonal antibodies
CC can also be used to detect and diagnose Crohn's disease.

XX
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 21; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLVQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLVQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNLSGLSLAATTSPQLLACGPTVHTQCSNTVYKGLCLFGLNLRQPOK 120
DB 77 RLQVPVEAVNLSGLSLAATTSPQLLACGPTVHTQCSNTVYKGLCLFGLNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLDGSIIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEF 180
DB 137 FPEALRGCPQEDSDIAFLDGSIIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEF 196

QY 181 RIHFYFKFQNNPRSLKPIQTLLGRTHTTATGLRKVVRELFNITNGARKNAKILFLL 240
DB 197 RIHFYFKFQNNPRSLKPIQTLLGRTHTTATGLRKVVRELFNITNGARKNAKILVVI 256

QY 241 TDGEKFGDPLGYEDVJPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
DB 257 TDGEKFGDPLGYEDVJPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316

QY 301 NFEALKTQVONOREKIFATEGTQSGSSFEHMSOEGFSAITNSGPLLSTVGSVDWAG 360
DB 317 NFEALKTQVONOREKIFATEGTQSGSSFEHMSOEGFSAITNSGPLLSTVGSVDWAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVNSGSDTLVLI GAPHYEQTRGGQSVSCPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVNSGSDTLVLI GAPHYEQTRGGQSVSCPL 496

QY 481 PRQQRARWOCDAVLYEQGPWGRFCAALTVLGDVNGDKLTDTVAIGAPGEEDNRGAVYLF 540
DB 497 PRQQRARWOCDAVLYEQGPWGRFCAALTVLGDVNGDKLTDTVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSORITAGSKLSPLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLLRQ 600
DB 557 HGTSGSGISPSHSORITAGSKLSPLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLLRQ 616

QY 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAVNETKNSTRQTQVLGLTQTCETILKQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAVNETKNSTRQTQVLGLTQTCETILKQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFAGNLRPVLAEDAQLFTALFPPEKNCNDNI CQDDLSITFSFMSLDCLVVG 780

DB 737 SLVGTPLSAFAGNLRPVLAEDAQLFTALFPPEKNCNDNI CQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFFFPDLDSYRKVSTLQNRORSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRQVTFFFPDLDSYRKVSTLQNRORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRSMQHOYQVSNLQBSLSISLVFLVPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRSMQHOYQVSNLQBSLSISLVFLVPV 976
QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1020
DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQET 1096
QY 1081 KVEPPEVDNPLPLIVGSSVGGILLILALITAAALYKLGFPKROYKDMMSGGPPGAEPO 1137
DB 1097 KVEPPEVDNPLPLIVGSSVGGILLILALITAAALYKLGFPKROYKDMMSGGPPGAEPO 1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX AC ABG61469;
DT 27-AUG-2002 (first entry)
XX DE Human Beta2 integrin alphaCD11b subunit.
KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX PN WO200230980-A2.
XX PD 18-APR-2002.
XX PF 15-OCT-2001; 2001WO-US32059.
XX PR 13-OCT-2000; 2000US-0688307.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX WPI; 2002-463260/49.
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
XX recovery, inhibiting locomotor damage, limiting locomotor impairment,
XX or limiting autonomic and sensory dysfunction following spinal cord
XX injury -
XX Example 5; Page 191-194; 270pp; English.
XX The invention relates to promoting locomotor recovery, inhibiting
XX locomotor damage, limiting locomotor impairment, or limiting autonomic
XX and sensory dysfunction following spinal cord injury by administering an

CC anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of
CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
CC molecule, vascular cell adhesion molecule). The method is useful for
CC promoting locomotor recovery, inhibiting locomotor damage, limiting
CC locomotor impairment, or limiting autonomic and sensory dysfunction
CC following spinal cord injury. In particular, the spinal cord injury
CC comprises compression of the spinal cord. The antibodies are also useful
CC for reducing inflammation at the site of a central nervous system injury.
CC The specification also details the identification of Beta2 integrin
CC alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2
CC integrins are implicated in diseases such as LAD (leukocyte adhesion
CC deficiency, inflammatory response, diabetes, multiple sclerosis,
CC arthritis, graft atherosclerosis, inflammatory bowel disease,
CC Crohn's disease, ulcerative colitis, immune complex alveolitis
CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
CC sequence included for comparison with the Beta2 integrin alpha protein
CC sequences.
XX
SQ Sequence 1153 AA;
Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTPPOLLAGCTVTHQTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTPPOLLAGCTVTHQTCSENTYVKGCLFLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDIAFLDGGSGIIIPHFRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
Db 137 PPEALRGCPQSDIAFLDGGSGIIIPHFRMKELVSTIMEQLKSKTFLSLMQYSEEF 196
Qy 181 RIHFTFEFQNNPRSLIKETITQLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 197 RIHFTFEFQNNPRSLIKETITQLGRTHATGLRKVVRLEFNITNGARKNAFKILV 256
Qy 241 TDGSKFGDPLGYEDVIELDREGVIRVYVGLGDAPFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGSKFGDPLGYEDVIELDREGVIRVYVGLGDAPFRSEKSRQELNTVASKPRDHVQAN 316
Qy 301 NFEALKTQVONLREKIFALECTGTGSSSFHEHMSQEGFSAATISNGPLLSVGSYDWAG 360
Db 317 NFEALKTQVONLREKIFALECTGTGSSSFHEHMSQEGFSAATISNGPLLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTREGQVSVCP 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTREGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGTSPSHSQRISAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLLRSQ 600
Db 557 HGTSGSGTSPSHSQRISAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRREGIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRREGIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRTROTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTROTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736

Qy 721 SLVGTPLSAFNGNLRPVLAEADAORLFTALPFPFKKNCNDNIQQDDLSTTFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGNLRPVLAEADAORLFTALPFPFKKNCNDNIQQDDLSTTFSFMSLDCLVVG 796
Qy 781 GRPEFNVTYVRNDGEDSVYRTQVTFEPLDLSYRKVSTLQNRORSQSWRLACESASSTEV 840
Db 797 GRPEFNVTYVRNDGEDSVYRTQVTFEPLDLSYRKVSTLQNRORSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANTVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANTVTSENMPRTNKTEF 916
Qy 901 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVLPV 960
Db 917 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVLPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORIOCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORIOCDIP 1036
Qy 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLILIVSTABILFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLILIVSTABILFNDVSFTLLPGQAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMMSSEGGPPGAEPQ 1153
RESULT 4
AAU80252
ID AAU80252 standard; Protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key
FT Misc-difference 499..500 /note="Encoded by GGG CAG AGG"
XX
PN WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US27227.
XX
PR 01-SEP-2000; 2000US-229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
XX
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilised in a desired conformation. The polypeptide of
CC the invention may have anti-inflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.
XX
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY '1 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVGAPQEIIVAANQKSLYOCYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVGAPQEIIVAANQKSLYOCYSTGSCPEI 76
QY 61 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNRQQPOK 120
DB 77 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNRQQPOK 136
QY 121 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTFKFQNNPRLSKPIPTQLLGRTHATGLRKVVRELFNITNGARKNAKILFL 240
DB 197 RIHFTFKFQNNPRLSKVPIPTQLLGRTHATGIRKVVRELFNITNGARKNAKILVVI 256
QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
DB 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316
QY 301 NFEALKTQVQUREKIFALEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTQVQUREKIFALEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVPLYTSKEKSTFINNTRVDSQNDAYLCYAAAILLRNRVQSVILGAPYQHIGLVAMER 420
DB 377 GVPLYTSKEKSTFINNTRVDSQNDAYLCYAAAILLRNRVQSVILGAPYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGYFASLCSDVDVNGSDTLVLI GAPHYEQTRGGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGYFASLCSDVDVNGSDTLVLI GAPHYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYILF 540
DB 497 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYILF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDITVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDITVGAQGHVLLRSQ 616
QY 601 PVLRVKAIIMEFPREVARNVFECNDQVKGKAGEVRVCLVHVKQSTRDLRREGQIQSVYT 660
DB 617 PVLRVKAIIMEFPREVARNVFECNDQVKGKAGEVRVCLVHVKQSTRDLRREGQIQSVYT 676
QY 661 YDLALDSGRPHGRAVNEFTKNSRRQTQVLGTLTQCTETIKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHGRAVNEFTKNSRRQTQVLGTLTQCTETIKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFAGNLRPVLAEADAQRFTALFPFEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEADAQRFTALFPFEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFPPFDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRQVTFPPFDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 856
QY 841 SGALKSTSCSNHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSNHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPKVAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPKVAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQSRSLPISLVFLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELRKAPVVCNCSIAVCQRIQCQIP 1020
DB 977 RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELRKAPVVCNCSIAVCQRIQCQIP 1036
QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFAVRSQTEF 1080
DB 1037 PFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFAVRSQTEF 1096
QY 1081 KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKQMMSEGGPPGAEPPQ 1137
DB 1097 KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKQMMSEGGPPGAEPPQ 1153
RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX AC AAO14428;
XX XX
XX 03-MAY-2002 (first entry)
XX DE Integrin Mac-1 alpha subunit.
XX KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
XX integrin related inflammatory disorder;
XX integrin related immunological disorder; rheumatoid arthritis; ischaemia;
XX reperfusion; hypovolemic shock; infarction; cerebral shock;
XX viral infection; cancer; gene therapy; vaccine;
XX bioactive agent screening.
XX OS Unidentified.
XX XX
XX WO200204521-A2.
XX PD 17-JAN-2002.
XX XX
XX 09-JUL-2001; 2001WO-US21805.
XX PF 07-JUL-2000; 2000US-216600P.
XX PR (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX PA (BLOO-) CENT BLOOD RES.
XX PI Springer T;
XX XX
XX WPI; 2002-148167/19.
XX PT New integrin I domain protein having alteration in at least 2
XX noncontiguous regions and exits in an open conformation, useful for
XX treating, preventing or suppressing inflammatory or immunological
XX disorders
XX PS Example 1; Fig 1F; 90pp; English.
XX XX
XX The invention comprises structurally biased variant integrin inserted (1)
XX domain proteins, wherein the alterations to the protein occur in at least
CC

CC two noncontinuous regions. Specifically the the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation,
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infection; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin.

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVAAVNRGSLVQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVAAVNRGSLVQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFGLFSGNLRRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFGLFSGNLRRQPOK 136

Qy 121 PPEALRGCPQSDIAFLIDGSGSIIPUDPRMKELVSTIMEQLKKSTLPFLMOYSEEP 180
Db 137 PPEALRGCPQSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKKSTLPFLMOYSEEP 196

Qy 181 RIHTTFKEFQNNPRSLKIPITOLGRTHPTATGLRKVVRELFINITGARKNAKFLFL 240
Db 197 RIHTTFKEFQNNPRSLKIPITOLGRTHPTATGLRKVVRELFINITGARKNAKFLV 256

Qy 241 TDGKFGDPLGYEDVPELDREGVIRYVGLFGDAFRSEKSKQELNTVASKPPRDHVFQAN 300
Db 257 TDGKFGDPLGYEDVPELDREGVIRYVGLFGDAFRSEKSKQELNTVASKPPRDHVFQVN 316

Qy 301 NFEALTKVQNLREKIPAEIGTQTGSSSPFHEMSQEGFSAATISNGPLLSVTGSDYDAG 360
Db 317 NFEALTKVQNLREKIPAEIGTQTGSSSPFHEMSQEGFSAATISNGPLLSVTGSDYDAG 376

Qy 361 GVFLYTSKEKSTFTNMTVDSDMDAYLGVAALILNRVOSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFTNMTVDSDMDAYLGVAALILNRVOSLVGLGAPRYOHIGLVAMFR 436

Qy 421 QNTGWESNANVKGITQICAYFGASLCSVDVDSNGSTDVLVLCAPHYVEQTRGGQVSVCP 480
Db 437 QNTGWESNANVKGITQICAYFGASLCSVDVDSNGSTDVLVLCAPHYVEQTRGGQVSVCP 496

Qy 481 PRGQARWQCDVLYGEGQPWGRFGAALTVLGVDVNGDKLTVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPWGRFGAALTVLGVDVNGDKLTVAIGAPGEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLMGLVDLTVAAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLMGLVDLTVAAGHVLRLRSQ 616

Qy 601 PVLVRKALMEFNPVARENVFECNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKALMEFNPVARENVFECNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDGRPHSRVFNTEKNSRQTVGLTQTCETLKLQLPNCIEDPVSPIVRLNLF 720
Db 677 YDLALDGRPHSRVFNTEKNSRQTVGLTQTCETLKLQLPNCIEDPVSPIVRLNLF 736

Qy 721 SLVGTPLSAFGNLRVLAEADQRLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRVLAEADQRLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 916
Qy 901 QLELPVKYAVYVMTSHGVSSTKYNFTASENTSRVVMQHVQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPVKYAVYVMTSHGVSSTKYNFTASENTSRVVMQHVQVSNLQORSLSPLSLVLPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1036

Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGOGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGOGAFVRSQTET 1096

Qy 1081 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPO 1137
Db 1097 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPO 1153

RESULT 6
AA04136
ID AA04136 standard; protein; 1153 AA.
XX AA04136;
AC AA04136;
DT 25-MAR-2003 (updated)
DT 07-SEP-1990 (first entry)
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
XX synthetic.

Key Location/Qualifiers
Modified-site 86..88 /*label= putative N-glycosylation site
Modified-site 240..242 /*label= putative N-glycosylation site
Modified-site 391..393 /*label= putative N-glycosylation site
Modified-site 469..471 /*label= putative N-glycosylation site
Modified-site 693..695 /*label= putative N-glycosylation site
Modified-site 697..699 /*label= putative N-glycosylation site
Modified-site 735..737 /*label= putative N-glycosylation site
Modified-site 802..804 /*label= putative N-glycosylation site
Modified-site 881..883 /*label= putative N-glycosylation site
Modified-site 901..903 /*label= putative N-glycosylation site
Modified-site 912..914 /*label= putative N-glycosylation site
Modified-site 941..943 /*label= putative N-glycosylation site
Modified-site 947..949 /*label= putative N-glycosylation site
Modified-site 979..981 /*label= putative N-glycosylation site
Modified-site 994..996 /*label= putative N-glycosylation site
Modified-site 1022..1024 /*label= putative N-glycosylation site

```

FT Modified-site 1045..1047
FT FT /*label= putative N-glycosylation site
FT Modified-site 1051..1053
FT FT /*label= putative N-glycosylation site
FT Modified-site 1076..1078
FT FT /*label= putative N-glycosylation site
FT region 1..16
FT FT /*label= signal peptide
FT region 1106..1134
FT FT /*label= putative transmembrane region
XX EP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-0115159.
XX
XX 23-AUG-1988; 88US-0235353.
XX 09-MAR-1989; 89US-0321239.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Springer TA, Corbi A;
XX
XX WPI; 1990-125938/17.
XX N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
XX inflammation and viral infections, and in diagnosis
XX
XX Disclosure; Page ?; ?pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
XX recognition of and migration to sites of inflammation. It also attaches
XX to cellular substrates as part of this function making it useful in
XX visualising endothelial tissue.
XX
XX Mac-1 is a member of the Integrin Gene superfamily.
XX
XX (Updated on 25-MAR-2003 to correct PR field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1153 AA;

Query Match 99.1%; Score 5829; DB 11; Length 1153;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1123; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEEF 196

QY 181 RIHFTKEFQNNPNRSLKPIITQLLORTHTATGLRKVVRELPIYNGARKNAFKILFL 240
Db 197 RIHFTKEFQNNPNRSLKPIITQLLORTHTATGLRKVVRELPIYNGARKNAFKILVVI 256

QY 241 TDGEKFGDPLGYEDVPELDREGVIRVILGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRVILGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 316

QY 301 NFPEALTKVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITSGPLLLSTVGSYDWAG 360
Db 317 NFPEALTKVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITSGPLLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLVGAAAIILRNVRQSVLVGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLVGAAAIILRNVRQSVLVGAPRYQHIGLVAMPR 436

```

```

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVSCPL 496
QY 481 PRQBARWCCDAVLIGEQQPGWRGGAALTIVLDVNGDKLTDVAIGAPGEENRGAVYLF 540
Db 497 PRQBARWCCDAVLIGEQQPGWRGGAALTIVLDVNGDKLTDVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVCQAQGHVLLLRQ 600
Db 557 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVCQAQGHVLLLRQ 616
QY 601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQIPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQIPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNGNLRPVLAEDAORLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGNLRPVLAEDAORLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSMWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSMWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVTYKLNFTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVTYKLNFTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036
QY 1021 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIIPNDVSVFTLLPQGGAFVRSQTET 1080
Db 1037 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIIPNDVSVFTLLPQGGAFVRSQTET 1096
QY 1081 KVEPEFVNPLLIYGVSSVGGLLLALITAAALYKLGFFKRYKDMWSEGGPGAEPO 1137
Db 1097 KVEPEFVNPLLIYGVSSVGGLLLALITAAALYKLGFFKRYKDMWSEGGPGAEPO 1153

RESULT 7
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (updated)
DT 05-FEB-1991 (first entry)
XX
XX p150.95 alpha subunit encoded by clone lambdaX47.
DE p150.95 leucocyte adhesion receptor alpha-subunit;
KW hairy cell leukaemia; rhinovirus.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
PH Modified-site 61..63
FT Modified-site /label=glycosylation site
FT Modified-site 89..91
FT Modified-site /label=glycosylation site
FT Modified-site 385..387
FT Modified-site /label=glycosylation site

```


OS Homo sapiens.
 XX WO200029446-A1.
 XX 25-MAY-2000.
 XX 16-NOV-1999; 99WO-US27139;
 XX 16-NOV-1998; 98US-0193043;
 PR 08-JUL-1999; 99US-0350259;
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 PI WPI; 2000-387751/33.
 XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous
 PT system injury sites
 XX Example 5; Fig 1; 270pp; English.
 PS Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit; the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This
 CC sequence was used in an alignment to identify a novel beta2 integrin
 CC alpha subunit; alpha d (AA060014 and AAB07359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha d. The
 CC Alpha d gene and protein may be useful in therapy for diseases linked
 CC to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis,
 CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
 CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
 CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
 CC antibodies may be used in the inhibition of macrophage infiltration at
 CC the site of a central nervous system injury. The monoclonal antibodies
 CC can also be used to detect and diagnose Crohn's disease.
 XX Sequence 1163 AA;
 SQ
 Query Match 58.2%; Score 3423; DB 21; Length 1163;
 Best Local Similarity 60.2%; Pred. No. 3.2e-274;
 Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVGVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 60
 DB 20 FNLDTEELTAFRVDSAGRGDSVVQVANSWVVGATQKIIAANQIGLYQCGYSTGACEPI 79
 QY 61 RLQVPVEAVNMSLGLSLAATSPOLLACGPVTHQTCSENTYVVKGLCFGLGNLRQOPQK 120
 DB 80 GLQVPPEAVNMSLGLSLASTTSPQLACGPVHHECGRNMYLTGLCFLLGPT--QLTQR 137
 QY 121 FEARLGCQPEQSDTAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSBEF 180
 DB 138 LPVSQECQPEQSDIVFLIDGSGSISSRNFAFMNFRAVISQFORPSTQFSLMQFSNKF 197
 QY 181 RIHTFKEFQNNPRLSKPTITOLLGRTHATGLRKVVRELFNITNGARKNAKIFLL 240
 DB 198 QTHFTFEEPRRTSNLSLASHVQLQGYTYTATQNVVRLPHASYGARRDAIKILIVI 257
 QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLFGDFAFRSEKSRQELNTVASKPPDRHVQAN 300
 DB 258 TDGKXGSDLDYKDVIPMAADAGIIRYAIGVGLAFQNRNSWELNDIASKPSQEHIFKVE 317
 QY 301 NFEALKTQNLQREKIFALETGTQSGSSSFEHMSQEGFSAITNSGPLLSTVGSYDNAG 360
 DB 318 DFDALKDIOQLQKEKIFALETGTETISSSSFELEMAQEGFSVFTPDGPVLGAVGFTWSG 377
 QY 361 GWFLYTSKEKSTFINNRVDSMDNDAYLGAAAILLRNRVQSLVILGAPRYQHIGLVAMFR 420

DB 378 GAFLYPPNMSPTFINNSQENVDNRDSYLGYSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
 QY 421 QNTGMVESNANVKGTOIGAYFGASLCSVDVDSNGSLDLVLIGAPHYYEOTRGQSVSCPL 480
 DB 438 QVSRQWRMKAIEVIGTOIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQSVSCPL 497
 QY 481 PRQORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYVLF 540
 DB 498 PRGWR-RWWCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVVGAPGEEENRGAYVLF 556
 QY 541 HGTSGSGISPSHSQRTAGSKLSPLQYFGQSLSGGODLTMDGLVDLTGVAGQHVLLRSQ 600
 DB 557 HGVLPSPISPSHSQRTAGSKLSPLQYFGQSLSGGODLTMDGLVDLTGVAGQHVLLRTR 616
 QY 601 PVLVRKAIMEFNPREVARNVFECDVWVGKEAGEVRVCLVHVKSTRDRREGQIQSVVT 660
 DB 617 PVLWGVSMQFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDVPSPILVRLNF 720
 DB 677 LDALAPGRLS PRAIFQETKNRSLSRVRLGLKARCENFLLPSCVEDSVIPIILRLNF 736
 QY 721 SLVGTPLSAGNLRPLVLAEDAQRLFTALFFPKXCGNDNICODDLSITFFMSLDCLVVG 780
 DB 737 TLVGKPLLAFLNLRPLMAALAQAQRYFTASLPFEKXCGADHICODNLGHSFSPGLKSLVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSQRWRLACESASTEV 840
 DB 797 SNLELNAEVMWNDGEDSYGTTTFSSHAPGLSRYVAEGQKQGLRSLHLTC--CSAPVG 854
 QY 841 SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
 DB 855 SQGTWSTFSCRINHLIFRGAQITFLATFDVSPRAVGLDRLLLIANVSSENIPRTSKTIF 914
 QY 901 QLELPVKYAVMYVTSHGVSTKYLNFATAS-ENTSRVMQHOYOVSNLQORSIPISLVLP 959
 DB 915 QLELPVKYAVYVIVSSHEQFTKYLNFSSEKESHVAMHYQVNNLQORLPSVINFWPV 974
 QY 960 VRLNQTWDRPOVTFSENLSSTCTKRLPSHSDFLAELRKAPVNCSTAVCQRIQCDI 1019
 DB 975 VELNQEAVMDVEVSHQPNFSLRCSSEKIAPPASDFLAHTQKNPVLQCSAGCLRFRCDV 1034
 QY 1020 PFFGIOBEFNATLKNLSFDWYIKTSHNHLILVSTAILFNDSVFTLLPGGAFVRQTE 1079
 DB 1035 PSFSVQEEELDFTLKGNLSFGWVRQILQKVSVSVSVAEIIPTDSVYSQLPQGEAFMRAQTI 1094
 QY 1080 TKVEPFEVPNPLPIVGVSSVGGLLLLALITAAALKGLGFFKRYQKDMSE 1128
 DB 1095 TVLEKYKVHNPILVIGSSIGGLLLALITAVLYKVGFFKRYQKEMME 1143
 RESULT 10
 ABG61470
 ID ABG61470 standard; Protein; 1163 AA.
 XX
 AC ABG61470;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human Beta2 integrin alphaCD11c subunit.
 XX
 KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
 KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 XX
 OS Homo sapiens.
 XX

PF 08-APR-2002; 2002WO-US10824.
XX 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
FA Sun Z, Jay G;
XX WPI; 2003-058520/05.
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
PS Claim 1; Page 225-228; 416pp; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer. e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer.

Query Match 58.2%; Score 3421; DB 24; Length 1163;
Best Local Similarity 60.1%; Pred. No. 4.7e-274;
Matches 679; Conservative 141; Mismatches 303; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGQSVVQVQGSRRVVVGAPOEIVAAQNRGLSYQCDYSTGSCBPI 60
DB 20 FNLDTEELTAFRVDSAGSDSVVQVANSVWVGAPOKITAANQTGGLYQCGYSTGACBPI 79

QY 61 RLQVPVAVNMGLSLAATSPQALLACGPTVHTCSENTYVKGCLFGSNLQPOK 120
DB 80 GLQVPVAVNMGLSLAATSPQALLACGPTVHTCGRNMVLTGLCFLGLGT--QLTOR 137

QY 121 FPEALRGCPQEDSDFLIDGSGSIIPHDFRMKELVSTIMEQLKKSKTLFSLMQYSBEF 180
DB 138 LPVSRQECRQDQIVFLIDGSGSISSRNFAFMNFVRAVISOQFQPSQFSLMQFSNKF 197

QY 181 RIHFTKPEQNPNRSLIKPITQLGRTHRTATGLRKVVRELFNITNGARKNAFKILFL 240
DB 182 RIHFTKPEQNPNRSLIKPITQLGRTHRTATGLRKVVRELFNITNGARKNAFKILFL 240

Db 198 QTHLTTEFRRTSNPLSLASVHQLQGYTTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDGEKFGDPLGYEDVIPLELDREGVIRYVLGFCDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 258 TDGKKEGDTLDYKDVIPMADAAGIIRYAGVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDMAG 360
Db 318 DFDALDKDIQQLREKIFPIEGTETSSSFELEMAQEGFSAVFTPOGVLGAVGSGFTWSG 377
QY 361 GVFLYTSKEKSTFINMTRVDSDMNDAYLGVAAYAIILNRVQSLVLGAPRYQHILGLVAMFR 420
Db 378 GARLYPPNMSPTFINNSQENVDMRDSYLGYSTELALWKGVSLVGLGAPRYQHTGKAVIFT 437
QY 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQVSVCP 480
Db 438 QVSRQWRMKAENVGTQIGSYFGPSLCSVDVDSGDTDLVLIGPHYEYQTRGAQVSVCP 497
QY 481 PRQARWQCDVLYGEOGPWCRFGAALTVDLVGNDKLTDTVAIGAPGEDNRGAYVLF 540
Db 498 PRQWR-RWMCDAVLYGEOGHPWCRFGAALTVDLVGNDKLTDTVAIGAPGEDNRGAYVLF 556
QY 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGOSLGGODLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGVLPISPSHSORISAGSKLSPRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLKVAIMEFNPREVARNVFECDNVQVKGKAGEVRVCLHVQKSTRDRREGOIQSVVT 660
Db 617 PVLWVGVMQFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSVT 676
QY 661 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVPVILRLNF 720
Db 677 LDALDPRLSRATFOETKRSLSRVRLGLKAHCENFLLPSCVEDSVTPITRLNF 736
QY 721 SLVGTPLSARGNLPLVLAEDAORLTALFPFKXNCNDNICODDLSITFFSMJDLCLVVG 780
Db 737 TLVGKPLAFLNPLMLAADAQRYFTASLPFKXNCADHICQDNLGISFSFPGLSLLVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 SNLELNAEVMVNDGEDSYGTTITFSHPAGLSRYVAEGQKQGLRSLHLTCDSPAVG-- 854
QY 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 855 SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914
QY 901 QLELPVKYAVVMVTSHGVTXYLNFTAS-ENTSRVMQHOYQVSNLQORSLSLFLVP 959
Db 915 QLELPVKYAVTVVSSHEQTKYLNFSSEKESHVAMHRVOVNNLQORDLPVSNFWP 974
QY 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDI 1019
Db 975 VELNQEAVMVDVEVSLPQNPSLRCSSEKIAGPASDFLAHQKPNVLDCSIAGCLURFCDV 1034
QY 1020 PFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSTVFTLLPGOGAFVRSQTE 1079
Db 1035 PSFSVQBELDFTLKNLSFGVVRQILQKRVSWVAEITFTDTSVYSQLPQGEAFMRAOTT 1094
QY 1080 TKVPEFPEVPNPLPLIVGSSVGLLLALLALITAAALYKLGFFKQYKDMWSE 1128
Db 1095 TVLEKYKVHNPPLIVGSSIGLLALLALITAVLYKVGFFKQYKEMWEE 1143

RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.

Homo sapiens.

Key Location/Qualifiers

Domain 17..1108

Region 150..352

/note= "extracellular domain"

/note= "this region is homologous to the insert common to CH11a,b,c and may be a site for interaction with ICAM family proteins"

Binding-site 465..474

/note= "putative cation binding site"

Binding-site 518..527

/note= "putative cation binding site"

Binding-site 592..600

/note= "putative cation binding site"

Region 1109..1128

/note= "transmembrane region"

Domain 1129..1161

/note= "cytoplasmic domain"

WO9517412-A1.

29-JUN-1995.

21-DEC-1994; 94WO-US14832.

05-AUG-1994; 94US-0286889.

23-DEC-1993; 93US-0173497.

(ICOS-) ICOS CORP.

Gallatin WM, Van Der Vieren M;

WPI; 1995-240603/31.

N-PSDB; AA091712.

Alpha' sub-unit polypeptide of human beta 2 integrin - used to

identify potential antiinflammatory agents, for the treatment of

graft arteriosclerosis, inflammatory bowel disease, asthma, etc.

Claim 7; Page 82-87; 172pp; English.

A probe based on a partial cDNA clone (given in AA091727) of canine

alpha-Tw1 was used to screen a human spleen cDNA library to identify

clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.

The cDNA was manipulated to allow expression of recombinant alpha-d

subunit in COS and CHO cells.

Sequence 1161 AA;

Query Match 57.6%; Score 3388; DB 16; Length 1161;

Best Local Similarity 58.9%; Pred. No. 2.6e-271;

Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

1 FNLDTENAMTFQENARGFGQSVVLOGSRVVGCAPOEIVAAANORGSLYOCDYSTGSCPEI 60

17 FNLDVEEPTIFQEDAGFGQSVVQFGSRLVWGAPLEVVAAANQTGRLYDCAATGMCQPI 76

61 RLQVPVEAVNMSLGLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGNLRQOPK 120

77 PLHIREAVNMSLGLTASTNGSRLLACGPTLHRVCGENSYSKSGCLLIGSRW-EIQT 135

121 FPEALRGCPQEDSIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKKSKTLFSLMYSSEF 180

136 VPDATPECPHOEMDIIVFLIDGSGSIDQNDFNQMKFQVAVMGQFECTDLFALMYSNLL 195

181 RIHFTKEQNPNPSLKPITQLLGRTHATGLRKVVRELPNITNGARKNAFKLFL 240

196 KIHFTTQRTSPSQOSLVDPIVLQKGLTFTATGILTVTQLPHHKNGARKSAKILLVI 255

KW	241	TDGEKFGDPLGYEDVIVPELDREGVIRYVGLFGDAPFRSEKSEQLNMTVASKPRDHVFOAN	300
KW	256	TDGQKYKDPLEYSVDVIPAOKERAGIIRYAIYGCHAPQGTARQELNTISSAPPQDHVFKVD	315
OS	301	NFEALKTYONOLREKIFAIEGTQTGSSSSFEHMSOEGFSAITSNGPLLSVTSYSDWAG	360
PH	316	NFAALGSIKOLQEKIYAVEGTQSRASSFOHMSOEGFSTALTMDDGLAVGAVGSWSG	375
FT	361	GVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR	420
FT	376	GAFLYPNMSPFTFINMSQENVMDRDSYLCYSTELALWKGVNLVGLGAPRYOHTKAVIFT	435
FT	421	QNTGMWESNANVKGTQIGAYFGASCSVDVDSNGSTDVLVIGAPHYYOTRGQGVSVCP	480
FT	436	QVSRQWRKKAQVETGTQIGSYFGASCSVDVDSNGSTDVLVIGAPHYYOTRGQGVSVCP	495
FT	481	PRGQBARWOCDAVLYGEOGPWGRGAALTVLGVGVNGDKLTDVAIGAGEEDNRCVAVLF	540
FT	496	PRGQRVQWOCDAVLRGEQHPWGRGAALTVLGVGVNGDKLTDVAIGAGEEDNRCVAVLF	555
FT	541	HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVBLTVGAQCHVLLRSQ	600
FT	556	HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLRSL	615
XX	601	PVLRYKAIEMFNPREVARNVFECDNVVYKGEAGEVVRVCLHVQKSTRDLREGIQSVVT	660
XX	616	PVLKVGAMRFPVEVAKAVYRCWBEKPSALEAGDATVCLTIQKSSDLQ--GDIQSSVR	673
XX	661	YDLALDSGRPHSRVAFNETKSTRQTQVGLGTQTCETLKLQLPNCIEDPVPVLRNLF	720
PR	674	FDLALDPGLTSLRAIFNETKPTLFRKTLGLGHICETLKLPLPCDVEDVVSPIILHNF	733
XX	721	SLVGTPLSAFGLNLRPVLAEADQRLFTALPPEKNGCNDNIQDDLSITFSFMSLCLVVG	780
XX	734	SLVREPIFSPQNLRPVLAVGSDQLFTASLPPEKNGCGQGLCEGLGVTLFSGLQTLTVG	793
XX	781	GPREFNVTVVRNDCGDSYRTQVTFPFPLDLISYRKVSTLQNRQSRWRLACESASSTEV	840
DR	794	SSLELNVIIVTNAGDSYGTIVSVLYPAGLSHRRVSGAOKPHQSAURLACETV-PTED	852
XX	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
XX	853	EG-LRSSRCVNHPIFHEGSGTFTVTFDVSYKATLGDRLMLRASSENKASSSKATF	911
PS	901	OLELPPKAVYVMVVTSHGVSTKYLNLF-TASENTSRVMOHOYOVSNLQSRSLPISLVFLVP	959
CC	912	QLELPVKAVYTMISRQSESTKYFNFAISDEKWKAEAEHRYKNNLSQDLAISINFWVP	971
CC	960	VLNQTIVWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI	1019
XX	972	VLLNGVAVWDVMEAPSQSL--PCVSEKRPQHSDFLTQISRSPMLDCSIADCLQRCOV	1029
QY	1020	PFPGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE	1079
DB	1030	PSFSVQOEBELDTLKNLSFGVWRETLOKKLVVSVAEITFDTSVYSQLPQGEAFMRAQME	1089
QY	1080	TKVEPPEFNPPLPLTVGSSVGLLALLITALYKLGFFKQYKQMMSE 1128	
DB	1090	MVLEDEEVNAPIIMSGSVGALLLALLITALYKLGFFKRYKHEMLED 1138	
RESULT	13		
AAW23049			
ID	AAW23049	standard; Protein; 1161 AA.	
XX	AAW23049;		
AC	AAW23049;		
XX	24-FEB-1998	(first entry)	
DT	24-FEB-1998		
XX	Human beta 2 integrin alpha d subunit.		
DE	Beta 2 integrin alpha d subunit; human; cell migration;		
XX	cell adhesion; phagocytosis; diabetes; atherosclerosis;		
KW			

multiple sclerosis; asthma; psoriasis; lung inflammation; acute respiratory distress syndrome; rheumatoid arthritis.

Homo sapiens.

Key	Location/Qualifiers
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1

Domain 17..1108

```
/label= Extracellular_domain
```

Domain	1109..1128
1109..1128	1109..1128

```
|
/label= Transmembrane_domain
/notes="homologous to the human protein"
```

/note= "homologous to the human CUL1 transmembrane region"

Domain 1129..116

```

domain
/label= Cytoplasmic domain

```

Domain 150..352

/note= "region homologous to the I (insertion)

WO9731099-A1.

28-AUG-1997.

24-FEB-1997: 97WO-US02713.

22-FEB-1996: 96US-0605672.

(ICOS-) ICOS CORP.

Gallatin WM. van Der vieren M:

WPI: 1997-435154/40.

N-PSDB; AAT79220.

Hybridoma 199M and antibody secreted by it - specific for new rat beta2 integrin subunit, useful to detect subunit in cells and modulate its activity

Example 5; Page 116-120; 222pp; English.

This polypeptide comprises a novel human beta 2 integrin subunit, designated alpha d. Its sequence was deduced from a cDNA clone (see AAT79220) isolated from a spleen cDNA library. Alpha d is involved in cell migration, phagocytosis and cell-cell interaction. Recombinant alpha d polypeptides can be expressed in transformed host cells for use in assays for identifying antibodies or other compounds that modulate alpha d activity or which modulate the interaction between alpha d and a ligand, for treating or preventing diseases in which macrophages are implicated. Treatment is applicable to disease states in which alpha d binding, or localised accumulation of cells which express alpha d, is implicated such as such as type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis.

Sequence 1161 AA:

every Match	57.6%	Score 3388:	DB 18:	Length 1161:
-------------	-------	-------------	--------	--------------

Best Local Similarity	58.9%	Pred. No. 2.6e-271;
-----------------------	-------	---------------------

atches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAAQRGSLYQCDYSTGSCEPI 60

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

121 FBFBY BCCBDBEBCDIAEITDCCSCCTDDBDEBDMVEI VCTIMEOI VVSVYI BCI MOVSEEE 180

136 VDDX TDECPHAE MIVET I DCSST DONDENOMKCEVIAIMCOEECTOTI EAT MOYENI I 195

XX

KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.

XX

OS Homo sapiens.

XX

PN US581029-A.

XX

PD 03-NOV-1998.

XX

PF 07-JUN-1995; 95US-0482293.

XX

PR 07-JUN-1995; 95US-0482293.

XX

PR 23-DEC-1993; 93US-0173497.

XX

PR 05-AUG-1994; 94US-0286889.

XX

PR 21-DEC-1994; 94US-0362652.

XX

PA (ICOS-) ICOS CORP.

XX

PI Gallatin WM, Van Der Vieren M;

XX

DR WPI; 1998-609318/51.

XX

DR N-PSDB; AAV67281.

XX

PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 PT immunohistochemical analysis

XX

PS Example 5; Column 61-66; 106pp; English.

XX

CC The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotype antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.

XX

SQ Sequence 1161 AA;

XX

Query Match 57.6%; Score 3388; DB 19; Length 1161;

XX

Best Local Similarity 58.9%; Pred. No. 2.6e-271;

XX

Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

QY

1 FNLDTENAMTFQENARGFGQSVVLOGSRVWVGAPQEIIVAANORGSLYOCXYSTGSCPI 60

DB

17 FNLDVEEPTIFQEDAGFGQSVVQFGSLVVGAPLEVAANQGTGLYDCAATGMCQPI 76

QY

61 RLQVPVEAVNMSLGLAATTPSPOLLACGPTVHOTCSENTYVKGCLFLFGSLNLRQOPQK 120

DB

77 PLHIRPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-ELIQT 135

QY

121 FPEALRGCEQDSDIAFLIDGSGSIIPDFRPMKELVSTIMEQLKSKTLFSLMQVSEBF 180

DB

136 VPDATPECPHQEMDVIPLIDGSGSIDQNDNFQMKGFVQAVMGQFEGDTLFFALMQVSNLL 195

QY

181 RIHFTKEFQNNPNPSLKPITQLLGRTHRTATGLRKVVRELPNITNGARKNAFKTLFL 240

DB

196 KIHFTIQTPTSPSQSLVDPIVOLKGLTFTATGILTVVTQFHHKNGARKSAKTLI 255

QY

241 TDGEKFGDPLGVEDVPIPELDREGVIRYVLGFCGDAFRSEKSRQELNTVASKPRDHDVFOAN 300

DB

256 TDGQKYKDPLEYSVDVPIQEKAGIIRYATGVGHAFQPTARQELNTISSAPQDHFVKVD 315

QY

301 NFEALKTQVQNKREKIFALIEGTQGTSSSFEHMSQEGFSAATSNGLPLSTVGSVDWAG 360

DB

316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLGLGAVGSPWSG 375

QY

361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAAIILNRRVQSLVLGAPRYQHIGLVAMPR 420

DB

376 GAFLPPNMSPTFINMSQENVDMRDSYLGYSYELALMKGVQNLVLGAPRYQHTKAVIFT 435

QY

421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQVSVCP 480

DB

436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDGSTDILIGAPHYYEOTRGQVSVCP 495

QY

481 PRGORARWOCDAVLVGEQOPMGRFGAALTVLGDVNGDKLTDVAIGACEEDNRCVAYLF 540

DB

496 PRGQRVQWOCDAVLVGEQOPMGRFGAALTVLGDVNEDKLTDVAIGACEQENRCVAYLF 555

QY

541 HGTSGSIGSPSHSQRISAGSKLSPRIQYFQCSLGGQDLTMGDLVLTVAQOQHVLRLSQ 600

DB

556 HGASESGISPSHSQRISAGSKLSPRIQYFQCSLGGQDLTMGDLMDLAVGARGQVLLRLSL 615

QY

601 PVLRVKATMEFNPVARNVFECDQVVKGKAGEVVRVCLHVQKSTRRLRREGQIQSVVT 660

DB

616 PVLKVGAMRFSPVEVAKAVYRCWEKPSALEAGDATVCLTTQKSLDQL--GDIQSSVR 673

QY

661 YDLALDSGRPHSRVFNKSTRQTQVLGLTCTETLKLQLPNCIEDPVPVILRLNF 720

DB

674 FDLALDPGLRLTSRAIFNETKNPTLTKRKTGLGHIHCETLKLPLPCVEDVDVSPILHLNF 733

QY

721 SLVGTPLSAFGNLRPLVLAEDAQRLEFALFPFPEKNGCNDNICQDDLSITFSFMSLDCLVVG 780

DB

734 SLVREPIPSQNLRPVLAVGSQDLFTASLPFEKNGQDGLCEGDLGVTLSFGLQTLTVG 793

QY

781 GPREFNVTVRNDEGDSYRTOVTFFPLDLSYKRVSTLQNRQSRWRLACESASSTEV 840

DB

794 SSLELNVTVMNAGEDSYGTVSVLYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852

QY

841 SCALKSTSCSINHPIFPENSEVTFTEDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

DB

853 EG-LRKSQCSVNHPIFHEGSGNGTFIVTFDVSYKATLGDRLMLRASSENKASSKATF 911

QY

901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVP 959

DB

912 QLELPVKYAVYTMISROBESTKYFNPAFSDKKKEAEHRYRVNLSQRLDAISINFWVP 971

QY

960 VRLNQTIVWRDQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIAVCQRIQCDI 1019

DB

972 VLLNGVAVMDVMEAPSQSL--PCVSEKPKPHQSDFLTQISRSPLMDCSIADCLQPRCDV 1029

QY

1020 PFGFQOEENATLKNLSFDVYIKTSHNHLIVSTAELFNDSVETLLPGCAFVRSOTE 1079

DB

1030 PSFSVQOEELDTLKNLSFGWVRRETLQKVLVSVVAEITFTDTSVYSQLPQGEAFMRAQME 1089

QY

1080 TKVEPFEPNPLPLIVGSSVGLLALLALITAAALYKLGFFKQYKQWMSSE 1128

DB

1090 MYLEDEYVNAIPIIMGSSVGALLLALITATLYKLGFFKRYKHEMLED 1138

XX

RESULT 15

XX

AAW65089

XX

ID AAW65089 standard; Protein; 1161 AA.

XX

AC AAW65089;

XX

XX AAW65089;

DT

28-SEP-1998 (first entry)

XX

Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.

DE

Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;

type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;

lung inflammation; acute respiratory distress syndrome;

rheumatoid arthritis.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Peptide 1..16

FT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:20:45 ; Search time 26.1265 Seconds
(without alignments)
8026.803 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFCQ.....FKRQKDMSEGGFCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5879	100.0	1137	11	US-09-902-481A-3
2	5857	99.6	1137	11	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	11	US-09-902-481A-1
5	5839	99.3	1153	11	US-09-891-943-3
6	5839	99.3	1153	15	US-10-144-259-30
7	5839	99.3	1153	15	US-10-207-655-176
8	5836	99.3	1137	11	US-09-902-481A-5
9	5832	99.2	1137	11	US-09-902-481A-6
10	5823.5	99.1	1152	10	US-09-945-265-4
11	3446	58.6	1163	12	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	11	US-09-891-943-4
14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	11	US-09-891-943-2

16	3372.5	57.4	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3372.5	57.4	1161	11	US-09-891-943-99	Sequence 99, Appl
18	3213.5	54.7	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3213.5	54.7	1161	11	US-09-891-943-55	Sequence 55, Appl
20	3201	54.4	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3192.5	54.3	1161	11	US-09-891-943-53	Sequence 53, Appl
22	3192.5	54.3	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3192.5	54.3	1151	11	US-09-891-943-37	Sequence 37, Appl
24	3180	54.1	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3180	54.1	1155	11	US-09-891-943-46	Sequence 46, Appl
26	1536.5	26.1	1170	10	US-09-945-265-2	Sequence 2, Appl
27	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, Appl
28	1229.5	20.9	494	11	US-09-891-943-103	Sequence 103, Appl
29	1128.5	19.2	413	9	US-09-350-259-101	Sequence 101, Appl
30	1128.5	19.2	413	11	US-09-891-943-101	Sequence 101, Appl
31	1128	19.2	1179	12	US-10-177-550-2	Sequence 2, Appl
32	1085.5	18.5	1151	11	US-09-984-130-103	Sequence 103, Appl
33	1085.5	18.5	1151	12	US-09-836-353A-103	Sequence 103, Appl
34	1079.5	18.4	1189	11	US-09-984-130-35	Sequence 35, Appl
35	1079.5	18.4	1189	12	US-09-836-353A-35	Sequence 35, Appl
36	1054	17.9	1181	15	US-10-160-354-2	Sequence 2, Appl
37	1004	17.1	216	12	US-09-795-872-5	Sequence 5, Appl
38	964	16.4	1034	11	US-09-984-130-43	Sequence 43, Appl
39	964	16.4	1034	12	US-09-836-353A-43	Sequence 43, Appl
40	950	16.2	240	15	US-10-102-806-453	Sequence 453, Appl
41	941	16.0	191	11	US-09-805-354-1	Sequence 1, Appl
42	941	16.0	191	15	US-10-144-259-1	Sequence 1, Appl
43	852	14.5	707	9	US-09-764-870-313	Sequence 313, Appl
44	852	14.5	707	15	US-10-125-540-313	Sequence 313, Appl
45	807	13.7	223	12	US-10-002-631C-197	Sequence 197, Appl

ALIGNMENTS

RESULT 1

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 11; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFCQSVVQLOGSRVVGAPQEIVAANQKGLGYCDYVSTGSCPEI	60
Db	1	FNLDTENAMTFQENARGFCQSVVQLOGSRVVGAPQEIVAANQKGLGYCDYVSTGSCPEI	60
Qy	61	RLOQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGICFLFGSNLRQOPK	120
Db	61	RLOQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGICFLFGSNLRQOPK	120
Qy	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFLSLMQYSEEF	180

```
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Qy 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPDRDHVFOAN 300
Db 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPDRDHVFOAN 300
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Qy 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Qy 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVSCPL 480
Db 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVSCPL 480
Qy 481 PRGORARWQCDVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 481 PRGORARWQCDVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
Db 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSPMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSPMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Qy 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960
Db 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960
Qy 961 RLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Db 961 RLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Qy 1021 PFGIQEEFNATLKGNSLSDWTIKTSHNHLIIVSTAELIIFNDSVFTLLPGQAFVRSQTE 1080
Db 1021 PFGIQEEFNATLKGNSLSDWTIKTSHNHLIIVSTAELIIFNDSVFTLLPGQAFVRSQTE 1080
Qy 1081 KVEPPEVNPPLIVGSSVGLLALITAAIYKLGFPKRYKDMSEGPPGAPQ 1137
Db 1081 KVEPPEVNPPLIVGSSVGLLALITAAIYKLGFPKRYKDMSEGPPGAPQ 1137
```

RESULT 2:

US-09-902-481A-4

; Sequence 4, Application US/09902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

```
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shiftman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4
```

```
Query Match 99.6%; Score 5857; DB 11; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVYKGLCFIFGSLNRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVYKGLCFIFGSLNRQOPQK 120
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Qy 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPDRDHVFOAN 300
Db 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPDRDHVFOAN 300
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Qy 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Qy 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVSCPL 480
Db 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVSCPL 480
Qy 481 PRGORARWQCDVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 481 PRGORARWQCDVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
Db 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSPMSLDCLVVG 780
```

Db 721 SLVGTPLSAFGLNLRPVLAEDAQRULTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 841 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 960
Db 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 960
QY 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1080
Db 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1080
QY 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGAEPO 1137
Db 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGAEPO 1137

RESULT 3

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQ 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQ 136
QY 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 196

RESULT 4

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia

QY 181 RIHFTKFEQNNPNRSLIKBITQLGRTHATGLRKVVRBELFNITNGARKNAPKILFL 240
Db 197 RIHFTKFEQNNPNRSLIKBITQLGRTHATGLRKVVRBELFNITNGARKNAPKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDEGVIRYVULGCDAPRSEKSRQELNTVASKPRDRHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVFQV 316
QY 301 NFEALKTQONOLREKI PAIEGTQTGSSSSPEHEMSQEGFSAATTSNGPLLSLTVGSYDNAG 360
Db 317 NFEALKTQONOLREKI PAIEGTQTGSSSSPEHEMSQEGFSAATTSNGPLLSLTVGSYDNAG 376
QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYQTGRTGGQVSVCP 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYQTGRTGGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVALF 540
Db 497 PRGQARWQCDVAVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVALF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGQDQDLMGDLVLTVGAQHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGQDQDLMGDLVLTVGAQHVLRLRSQ 616
QY 601 PVLRVKAIEMENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDRRLREGOIQSVVT 660
Db 617 PVLRVKAIEMENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDRRLREGOIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQPCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQPCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRULTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRULTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQORSQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 960
Db 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 976
QY 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSOTET 1096
QY 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGAEPO 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGAEPO 1153

APPLICANT: Mayo, Stephen
TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

FILE REFERENCE: A-70586-1/RT/RMS/RMK

CURRENT APPLICATION NUMBER: US/09/902,481A

CURRENT FILING DATE: 2001-07-09

PRIOR FILING DATE: 2000-07-07

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1153

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: mat_peptide

LOCATION: (17)..()

OTHER INFORMATION:

US-09-902-481A-1

Query Match 99.3%; Score 5839; DB 11; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQRGSLVQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQRGSLVQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCGTTHVQTCSENTRYVKGCLFGLGSLNRQOPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCGTTHVQTCSENTRYVKGCLFGLGSLNRQOPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHTTFKEFQNNPRSLVKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVVI 240
Db 197 RIHTTFKEFQNNPRSLVKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVIPEDREGVIRYVGLGDAFRSEKSKQELNTVASKPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVIPEDREGVIRYVGLGDAFRSEKSKQELNTVASKPRDHVFOAN 316
Qy 301 NFEALTKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 360
Db 317 NFEALTKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 480
Db 437 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 496
Qy 481 PRGQARWQCDAVLVGEQOPMGREFGAALTVLGDVNGDKLTDAITGAPGEDNRGAVLYF 540
Db 497 PRGQARWQCDAVLVGEQOPMGREFGAALTVLGDVNGDKLTDAITGAPGEDNRGAVLYF 556
Qy 541 HGTSGSGISPSHSQRISAGSLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRISAGSLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRLKATMEFNPVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVYT 660
Db 617 PVLRLKATMEFNPVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVYT 676
Qy 661 YDLALDGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
Qy 721 SLVGTPLSAFGNLRPVLAEADAQRFLTAFFPEKNCNGNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGNLRPVLAEADAQRFLTAFFPEKNCNGNDNICODDLSITFSFMSLDCLVWG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVKYSTLQONORSQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVKYSTLQONORSQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVNSLQORSIPISLVLVPV 960
Db 917 QLELPVKYAYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVNSLQORSIPISLVLVPV 976
Qy 961 RLNOTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNOTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDVSFTLLPGQGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDVSFTLLPGQGAFVRSQTET 1096
Qy 1081 KVEPPEVNPDLPLIVGSSVCGLLLLALITAAALKLGFKKQYKDMHSEGGPPGAEPQ 1137
Db 1097 KVEPPEVNPDLPLIVGSSVCGLLLLALITAAALKLGFKKQYKDMHSEGGPPGAEPQ 1153

RESULT 5

US-09-891-943-3

Sequence 3, Application US/09891943

Publication No. US20030077278A1

GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.

APPLICANT: Van der Vieren, Monica

TITLE OF INVENTION: No. US20030077278A1el Human 2

FILE REFERENCE: 27866/35004

CURRENT APPLICATION NUMBER: US/09/891,943

CURRENT FILING DATE: 2001-06-26

PRIOR APPLICATION NUMBER: 09/193,043

PRIOR FILING DATE: 1998-11-16

PRIOR APPLICATION NUMBER: 08/286,889

PRIOR FILING DATE: 1994-08-05

PRIOR APPLICATION NUMBER: 08/362,652

PRIOR FILING DATE: 1994-12-21

PRIOR APPLICATION NUMBER: 08/943,363

PRIOR FILING DATE: 1997-10-03

NUMBER OF SEQ ID NOS: 114

SOFTWARE: PatentIn ver. 2.0

SEQ ID NO 3

LENGTH: 1153

TYPE: PRT

ORGANISM: Homo sapiens

US-09-891-943-3

Query Match 99.3%; Score 5839; DB 11; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQRGSLVQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQRGSLVQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCGTTHVQTCSENTRYVKGCLFGLGSLNRQOPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCGTTHVQTCSENTRYVKGCLFGLGSLNRQOPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHTTFKEFQNNPRSLVKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVVI 240
Db 197 RIHTTFKEFQNNPRSLVKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVVI 256

```
QY 241 TDGEKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQBLNTVASKPRPDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVVLGFGDAFRSEKSRQBLNTVASKPRPDHVFQV 316
QY 301 NFALKTIVQNLREKIFAIEGTOTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 360
Db 317 NFALKTIVQNLREKIFAIEGTOTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSQNDAYLGVAALILRNVRQSLVGLGAPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSQNDAYLGVAALILRNVRQSLVGLGAPYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496
QY 481 PRQQRARWOCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWOCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVNETKNSRRQTQVLGTLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRAVNETKNSRRQTQVLGTLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFLTALPFPEKNCNDNICODDLSTIFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFLTALPFPEKNCNDNICODDLSTIFSMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKYSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 916
QY 901 QLELPVKYAYVMVTSHGYSKYLNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAYVMVTSHGYSKYLNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTET 1096
QY 1081 KVEPFEVPNPLPIVGVSSVGLLLALLIITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
Db 1097 KVEPFEVPNPLPIVGVSSVGLLLALLIITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153
```

RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
```

```
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-259-30
```

Query Match 99.3%; Score 5839; DB 15; Length 1153;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEI VAAANQORGLSYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEI VAAANQORGLSYQCDYSTGSCPEI 76
QY 61 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLFGSNLROPOK 120
Db 77 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLFGSNLROPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKRVVRELFININGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKRVVRELFININGARKNAFKILFLL 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQBLNTVASKPRPDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQBLNTVASKPRPDHVFQAN 316
QY 301 NFEALKTIVQNLREKIFAIEGTOTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 360
Db 317 NFEALKTIVQNLREKIFAIEGTOTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSQNDAYLGVAALILRNVRQSLVGLGAPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSQNDAYLGVAALILRNVRQSLVGLGAPYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496
QY 481 PRQQRARWOCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWOCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVNETKNSRRQTQVLGTLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRAVNETKNSRRQTQVLGTLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFLTALPFPEKNCNDNICODDLSTIFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFLTALPFPEKNCNDNICODDLSTIFSMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKYSTLQNRQSRWRLACESASSTEV 856
```

QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQVQVSNLQORSPLISLFLVPV 960
DB 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQVQVSNLQORSPLISLFLVPV 976
QY 961 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
DB 977 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036
QY 1021 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSOTET 1096
QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
DB 1097 KVEPFEPVNPPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207.655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 15; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVWOLQGSRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVWOLQGSRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLCFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQSDSIAFLIDGSGSTIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEP 180
DB 137 FPEALRGCPQSDSIAFLIDGSGSTIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEP 196
QY 181 RIHTFKEFQNNPRSLIKBITQLGRTHATGLRKVRELENTGARKNAKILFL 240
DB 197 RIHTFKEFQNNPRSLIKBITQLGRTHATGLRKVRELENTGARKNAKILVVI 256
QY 241 TDGEKFGDPLGVEDYVPELDEGVIRYVVGDAFRSEKSKQELNTVASKPRDRHVFQAN 300
DB 257 TDGEKFGDPLGVEDYVPELDEGVIRYVVGDAFRSEKSKQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTQNLREKIPAEIQTGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTQNLREKIPAEIQTGTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480

DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVIGAPGEENRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFCQSLSGGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFCQSLSGGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 660
DB 617 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEBPVSPVLRNLF 720
DB 677 YDLALDSGPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEBPVSPVLRNLF 736
QY 721 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACSSASTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACSSASTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQVQVSNLQORSPLISLFLVPV 960
DB 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQVQVSNLQORSPLISLFLVPV 976
QY 961 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
DB 977 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036
QY 1021 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSOTET 1096
QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
DB 1097 KVEPFEPVNPPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-11/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 99.3%; Score 5836; DB 11; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMSLGLSLAATTGPPQLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 120
DB 61 RLQVPVEAVNMSLGLSLAATTGPPQLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180

QY 181 RHFTFKFQNNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240
DB 181 RHFTFKFQNNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 300

QY 301 NFEALKTQONLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQONLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQTRGGQSVSCPL 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQTRGGQSVSCPL 480

QY 481 PRGQARWQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQHVLLRSQ 600
DB 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQHVLLRSQ 600

QY 601 PVLVRKAIINEFNPREVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLVRKAIINEFNPREVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660

QY 661 YDLALDSGRPHSRAVENETKNSTRQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRAVENETKNSTRQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720

QY 721 SLVGTPLSAFNLRPVLAEDAORLFTALPFFKNCNDNICODDLSITFSFMSLCLVVG 780
DB 721 SLVGTPLSAFNLRPVLAEDAORLFTALPFFKNCNDNICODDLSITFSFMSLCLVVG 780

QY 781 GPREFNVTVVRNDEGDSVRTQVTFEPFLDLSYRKVSTLQONRSORSWLACESASSTEV 840
DB 781 GPREFNVTVVRNDEGDSVRTQVTFEPFLDLSYRKVSTLQONRSORSWLACESASSTEV 840

QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

QY 901 QLELPVKYAVMVVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 960
DB 901 QLELPVKYAVMVVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLFLVPV 960

QY 961 RLNQTVIWRPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRQCDDIP 1020
DB 961 RLNQTVIWRPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRQCDDIP 1020

QY 1021 PFGIOEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPQGGAFVRSQTB 1080

DB 1021 PFGIOEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPQGGAFVRSQTB 1080

QY 1081 KVEPEVEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMWSEGGPPGABPO 1137
DB 1081 KVEPEVEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMWSEGGPPGABPO 1137

RESULT 9
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RWK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 11; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMSLGLSLAATTGPPQLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 120
DB 61 RLQVPVEAVNMSLGLSLAATTGPPQLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180

QY 181 RHFTFKFQNNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240
DB 181 RHFTFKFQNNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPRDHVFOAN 300

QY 301 NFEALKTQONLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQONLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQTRGGQSVSCPL 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQTRGGQSVSCPL 480

QY 481 PRGQARWQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

```
QY 541 HGTSGSGISPSHSORIASKLSRLOVFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSORIASKLSRLOVFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
QY 601 PVLRVKAIMFNFPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRREGQIOSVVT 660
DB 601 PVLRVKAIMFNFPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRREGQIOSVVT 660
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCLIEDVSPVLRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCLIEDVSPVLRLNF 720
QY 721 SLVGTPLSAFQNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFQNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRSLACSSASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRSLACSSASSTEV 840
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
QY 901 QLELPVKYAVYVMVTSYKLYNFTASENTSRVMOHQVSNLQORSIPSLVFLVPV 960
DB 901 QLELPVKYAVYVMVTSYKLYNFTASENTSRVMOHQVSNLQORSIPSLVFLVPV 960
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
QY 1021 FPGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1021 FPGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
QY 1081 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEQ 1137
DB 1081 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945.265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.1%; Score 5823.5; DB 10; Length 1152;
Best Local Similarity 98.8%;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 FNLDTENAMTQENARGQSVVQLQGSRVVVGAPQEIIVAAQNRCSLYQCDYSTGSCCEPI 60
DB 17 FNLDTENAMTQENARGQSVVQLQGSRVVVGAPQEIIVAAQNRCSLYQCDYSTGSCCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOK 120
```

RESULT 11

```
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOK 136
QY 121 FPEALRGCCPOESDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMQVSEEF 180
DB 137 FPEALRGCCPOESDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMQVSEEF 196
QY 181 RHFTFKFQNNPNRPSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RHFTFKFQNNPNRPSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILV 256
QY 241 TGEKFGDPLGVEDVJPELDREGVIRYVLFQGDAPFSEKSRQELNLTVAASKPRPDHVFQAN 300
DB 257 TGEKFGDPLGVEDVJPELDREGVIRYVLCVGDAPFSEKSRQELNLTVAASKPRPDHVFQAN 316
QY 301 NFEALTVQNLREKIFAIEGTQSGSSFEHMSQEGFSAATTSNGPLLLSTVGSVDWAG 360
DB 317 NFEALTVQNLREKIFAIEGTQSGSSFEHMSQEGFSAATTSNGPLLLSTVGSVDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILRNVRQSLVLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILRNVRQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 496
QY 481 PRGORARWOCDAVLVGEQGPWGRFCAALTVLGDVNGDKLTQVAGAPGEDNRGAYVLF 540
DB 497 PRG-RARWOCDAVLVGEQGPWGRFCAALTVLGDVNGDKLTQVAGAPGEDNRGAYVLF 555
QY 541 HGTSGSGISPSHSORIASKLSRLOVFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSORIASKLSRLOVFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 615
QY 601 PVLRVKAIMFNFPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRREGQIOSVVT 660
DB 616 PVLRVKAIMFNFPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRREGQIOSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCLIEDVSPVLRLNF 720
DB 676 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCLIEDVSPVLRLNF 735
QY 721 SLVGTPLSAFQNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFQNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRSLACSSASSTEV 840
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRSLACSSASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVYVMVTSYKLYNFTASENTSRVMOHQVSNLQORSIPSLVFLVPV 960
DB 916 QLELPVKYAVYVMVTSYKLYNFTASENTSRVMOHQVSNLQORSIPSLVFLVPV 975
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 976 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
QY 1021 FPGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1036 FPGIQEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1095
QY 1081 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEQ 1137
DB 1096 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEQ 1152
```



```
Db 138 LPVSRQCPROEQDIIFLIDSGSISRNFAFMVRAVISQFQRPSTQFSLMQFSNKP 197
Qy 181 RIHFTFEPQNNPRLIPIITQLQRTHTATGLRKVKVRELNITNGARKNAFKILFLL 240
Db 198 QTHFTFEPFRRTSNPLSLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKRSOELNTVASKPPRDHVFOAN 300
Db 258 TDGKKGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASPSGEHFKVE 317
Qy 301 NFEALKTQVQUREKIFAIEGTQTSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVMDRDSYLGYSTELALWKGVSQSLVGLGAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSCPL 480
Db 438 QVSRQWRMAKEVIGTQIGSVFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGGOVSCPL 497
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVDVNGDKLTDVVIAGPGEENRGAYVLF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGSLSGGQDLTMDGLVDLTGAGQHVLLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVLAAGARGQVLLLRTR 616
Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLVHOKSTRDRREGOIOSVVT 660
Db 617 PVLWVGVSQMFIPAEIPRSAFECEQVVSQTLVQSNICILYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCTIEDVPSVILRLNF 720
Db 677 LDLALAPGRLSPRAIFQETKNRSLSRVVLGLKAHCENFNLPLSPSCVEDSVIPIILRLNF 736
Qy 721 SLVQTPUSAFQNLARPVAEDAQRLFTALFPFEKCKGNNDTCQDDLSITFSPMSLDCLVVG 780
Db 737 TLVGKPLLAFLNRLPMLAALAAQRYFTASLPFEKCKGADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPREFNVTVRNDEGDSYRTQVTFPFLDLSYKVTSLQNRQSRMSRLACESASTEV 840
Db 797 SNLEINAEVMMNDGEDSYGTTITFSPAGLSYRYVAEGQKQGLRSLHUTC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVGLDRULLLIANVSSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTAS-ENTSRVMQHVQVSNLQORSPLISLFLVP 959
Db 915 QLELPVKYAVYVYVSSHEQTKYLNFSSEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
Qy 960 VRLNQTVMRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
Db 975 VELNQAEMMDVEVSHQNPNSLRCSSEKIAPPASDFLAHQKNPVLDCSTAGCLRFRCDV 1034
Qy 1020 PFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDSVFTLLPGOGAFVRQTE 1079
Db 1035 PSFSVQELDFTLKGNLSFGHVRQILQKVSVSVAEIIIFTSVYSQALPGQAEPMRAQTI 1094
Qy 1080 TKVBFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSE 1128
Db 1095 TVLEKYKVHNPILIVGSSIGCLLLALITAVLYKVGFFKRYQKEMEE 1143
```

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007728A1

; GENERAL INFORMATION:

```
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 1163
; ORGANISM: Homo sapiens
; US-09-891-943-4
```

Query Match 58.2%; Score 3423; DB 11; Length 1163;

Best Local Similarity 60.2%; Pred. No. 3e-307;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

```
Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLYOCYDSTGSCPEI 60
Db 20 FNLDTBELTAFRVDVAGFGDSVVQYANVVVGAPOKIITAAQIQGLYOCYSTGACEPI 79
Qy 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVYKGLCFGLGSLNRQPOQK 120
Db 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMVLTGLCFLLGPT--QLTQR 137
Qy 121 FPEALRGCEQSDSDIAFLIDGSGSIIPHPFRMKELVSTIMEOLKSKLTFSLMQVSEEF 180
Db 138 LPVSRQCPROEQDIIFLIDGSGSISRNFAFMVRAVISQFQRPSTQFSLMQFSNKP 197
Qy 181 RIHFTFEPQNNPRLIPIITQLQRTHTATGLRKVKVRELNITNGARKNAFKILFLL 240
Db 198 QTHFTFEPFRRTSNPLSLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKRSOELNTVASKPPRDHVFOAN 300
Db 258 TDGKKGDSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASPSGEHFKVE 317
Qy 301 NFEALKTQVQUREKIFAIEGTQTSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVMDRDSYLGYSTELALWKGVSQSLVGLGAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSCPL 480
Db 438 QVSRQWRMAKEVIGTQIGSVFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGGOVSCPL 497
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVDVNGDKLTDVVIAGPGEENRGAYVLF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGSLSGGQDLTMDGLVDLTGAGQHVLLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVLAAGARGQVLLLRTR 616
Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLVHOKSTRDRREGOIOSVVT 660
Db 617 PVLWVGVSQMFIPAEIPRSAFECEQVVSQTLVQSNICILYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCTIEDVPSVILRLNF 720
Db 677 LDLALAPGRLSPRAIFQETKNRSLSRVVLGLKAHCENFNLPLSPSCVEDSVIPIILRLNF 736
```


APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2003007278A1e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 2
LENGTH: 1161
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 57.6%; Score 3388; DB 11; Length 1161;
Best Local Similarity 58.9%; Pred. No. 5.3e-304;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

```
Qy 1 FNLDTENAMTQENARGEGSQVQLOGRVVGAPQEIIVAAANORGSLYQCDSYTGSCPEI 60
Db 17 FNLDOVEETIFQEDAGGFGQSVQFGSRLVVGAPLEVAAANQGRLYDCAATGMCPPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATSPOLLACGTPVHTQCSNTYVKGCLGFLGSGNLRQOPQK 120
Db 77 PLHTRPEAVNMSLGLTAASTNGSRLACGTPLHRVCGENSYSKGSCLLGSRW-ELIQT 135

Qy 121 FPEALRGCPQSDSDIAFLIDGSGIIPHDPRMKELVSTIMEQKSKTLFSLMQYSEEF 180
Db 136 VPDATPECPHQMEDIIVFLIDGSGIDQDNFQMKGFQVQAVMGQFEGDTLFLALMQYSNLL 195

Qy 181 RIHTEKFEQNNPNSRLIKPITOLLGRTHATGLRVKVVRELFNITNGARKNAFKILFL 240
Db 196 KIHFTFTQRTSPQSSLDVPIVQLKGLTFTATGLTIVTQLFHHKNGARKSAKKILIVI 255

Qy 241 TDGEKFGDPLGYEDYIPELDREGVIRYVVGFDAPRSEKQELNVTASKPPRDHVFOAN 300
Db 256 TDGQKYKDPLEYSVDVIPAERAGIIRYAIVGHAFOGPTARQELNTISSAPPQDHFVKVD 315

Qy 301 NFEALKTVQNLREKIFAIEGTQCTSSSFHEHMSQEGFSAITNSGPELLSTVGSYDWAG 360
Db 316 NFAALGSIQKQLOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFWSG 375

Qy 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAIIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 376 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELAWKGVQNLVLGAPRYOHTKAVIFT 435

Qy 421 QNTGWESNANVKGITGAYFGASICSDVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 436 QVSRQWRKKAETVGTIGSYFGASICSDVDVDSGSTDLLIGAPHYYEQTRGGQVSVCP 495

Qy 481 PRQBARWQCDVAVLYGEGQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEDNRCANVLF 540
Db 496 PRQQRVQWQCDVAVLRGEQHPMGRFGAALTVLGDVNEKLIIDVAIGAPCEQENRGAVYLF 555

Qy 541 HCTSGSISPSHSORIASKLSPLQYFGOSLGGODLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGASESGISPSHSORIASQSPLQYFGQALSGGODLTQDGLMDLAVGARGQVLLRLSL 615

Qy 601 PVLRYKAIENFNPVARNVPCNDQVVKKEAGEVRVCLHVQSKTRDRRLREGQIQSVVT 660
Db 616 PVLKVGAVMRFSPVEKAVYRCWEKPSALEAGATVCLTIQKSSLDQL--GDIQSSVR 673

Qy 661 YDLALDSGRPHSRVFNFTKNSRRQTQVLGTQTCETLKLQLPNCIEDPSPVIRLNF 720
Db 674 FDLALDPGRLTSLRAIFNETKNTPLTRKTLGLGHCETLKLALLPCDVEDVSPVIRLNF 733
```

```
Qy 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALPFFFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFFFEKNCGQDGLCEGLDGLVTLFSGLQTLTVG 793

Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSVRKYUSTLQNRQSQRSWRLACESASSTEV 840
Db 794 SSLELNVITVMNAGEDSYGTVVSLYPAGLSHRRVSGAQKQPHQSGALRLACETV-PTED 852

Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKLVNTSENNMPTNKTEF 900
Db 853 EG-LRSRCSVNHPIFHEGSGNGTFIVTFDVSYKATLGDRLMRASASSENKASSSKATF 911

Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMHQYQVSNLQGRSLPLISLVLVP 959
Db 912 QLELPVKYAVYTMISROESTKYFNATSDKKMKKEAHRVYVNNLSQRDLAISINFWVP 971

Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019
Db 972 VLLNGVAVMDVVMNEAPSQSL--PCVBERKPPQHSDFLTQISRSPMLDCSIADCLQPRCDV 1029

Qy 1020 PFFGIOEEFNATLKGNLSDWYIKTSHNHLLIVSTAEIILFNDSVFTLLPQCGAFVRSQTE 1079
Db 1030 PFSVQEEELDTLKGNLSCFWVRETLOKVLVSVVAEITPDTSVYQSLPQCEAFMRAQME 1089

Qy 1080 TKYEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFPKRYQKMMSE 1128
Db 1090 MVLEDEEVYNAIPIIMGSSVYVALLLALITATLTYKLGFFPKRYKEMLED 1138
```

Search completed: November 25, 2003, 14:41:11
Job time : 31.1265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using ew model

Run on: November 25, 2003, 14:13:09 ; Search time 13.5618 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKRQYKDMSEGGPGGAPBPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6C_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1	US-08-173-497-3
2	5839	99.3	1153	1	Sequence 3, Appli
3	5839	99.3	1153	1	Sequence 3, Appli
4	5839	99.3	1153	1	Sequence 3, Appli
5	5839	99.3	1153	2	Sequence 3, Appli
6	5839	99.3	1153	2	Sequence 3, Appli
7	5839	99.3	1153	2	Sequence 3, Appli
8	5839	99.3	1153	3	Sequence 3, Appli
9	5839	99.3	1153	4	Sequence 3, Appli
10	5808.5	98.8	1152	2	Sequence 43, Appl
11	5808.5	98.8	1152	5	Sequence 43, Appl
12	5808.5	98.8	1152	6	Patent No. 5424399
13	3446	58.6	1163	2	Sequence 44, Appl
14	3446	58.6	1163	5	Sequence 44, Appl
15	3423	58.2	1163	1	Sequence 4, Appli
16	3423	58.2	1163	1	Sequence 4, Appli
17	3423	58.2	1163	1	Sequence 4, Appli
18	3423	58.2	1163	1	Sequence 4, Appli
19	3423	58.2	1163	2	Sequence 4, Appli
20	3423	58.2	1163	2	Sequence 4, Appli
21	3423	58.2	1163	2	Sequence 4, Appli
22	3423	58.2	1163	3	Sequence 4, Appli
23	3423	58.2	1163	4	Sequence 4, Appli
24	3388	57.6	1161	1	Sequence 2, Appli
25	3388	57.6	1161	1	Sequence 2, Appli
26	3388	57.6	1161	1	Sequence 2, Appli
27	3388	57.6	1161	1	Sequence 2, Appli

28 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli
29 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli
30 3388 57.6 1161 3 US-08-943-363-2 Sequence 2, Appli
31 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli
32 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli
33 3372.5 57.4 1161 1 US-08-485-618-99 Sequence 99, Appl
34 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appl
35 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appl
36 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appl
37 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appl
38 3372.5 57.4 1161 4 US-09-688-307A-99 Sequence 99, Appl
39 3213.5 54.7 1161 3 US-09-193-043-55 Sequence 55, Appl
40 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appl
41 3206.5 54.5 1161 1 US-08-485-618-55 Sequence 55, Appl
42 3206.5 54.5 1161 1 US-08-362-652-55 Sequence 55, Appl
43 3206.5 54.5 1161 2 US-08-605-672-55 Sequence 55, Appl
44 3206.5 54.5 1161 2 US-08-482-293A-55 Sequence 55, Appl
45 3206.5 54.5 1161 2 US-08-943-363-55 Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEI VAAHQGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEI VAAHQGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 136
QY 121 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTFLFSLMOYSEEP 180
Db 137 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTFLFSLMOYSEEP 196
QY 181 RHFTFKFQNNPNRSLIKIPITQLGRTHATGLRKVRKVFNITGARKNAFKILFLL 240
Db 197 RHFTFKFQNNPNRSLIKIPITQLGRTHATGLRKVRKVFNITGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTQNLREKIFAIETGTQGTSSSSFEHMSQEGFSAATISNGPILLSSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIETGTQGTSSSSFEHMSQEGFSAATISNGPILLSSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLGAPHYETQTRGGVSVCP 480
Db 437 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLGAPHYETQTRGGVSVCP 496
QY 481 PRGARARQCDVAVLGEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGARARQCDVAVLGEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGSPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGAGCHVILLRSQ 600
Db 557 HGTSGSGSPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGAGCHVILLRSQ 616
QY 601 PVLRYKATMEFNPREVAARVNECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
Db 617 PVLRYKATMEFNPREVAARVNECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 676
QY 661 YDLALDGRPHSRAVFNTEKSTRQTVLGLTQTCETLKLQPNCTIEDPVSPIVRLNF 720
Db 677 YDLALDGRPHSRAVFNTEKSTRQTVLGLTQTCETLKLQPNCTIEDPVSPIVRLNF 736
QY 721 SLVGTPLSAFGLNRLPVLAEADQRLFTALPPEKNCNGNDNICQDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRLPVLAEADQRLFTALPPEKNCNGNDNICQDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDCGDSYRQVTFEFPDLISYRKVSTLQNRQSRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDCGDSYRQVTFEFPDLISYRKVSTLQNRQSRWSRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSYGVSTKYLNFNTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSYGVSTKYLNFNTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKPVNCNSIACVQICQDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKPVNCNSIACVQICQDIP 1036
QY 1021 PFGIQEEFNATLKGNSLDWYIKTSHNHLILVSTAEILFNDSVFTLLPGQAGAFVRSQTE 1080
Db 1037 PFGIQEEFNATLKGNSLDWYIKTSHNHLILVSTAEILFNDSVFTLLPGQAGAFVRSQTE 1096
QY 1081 KVEPEFVNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEFVNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 2
US-08-286-889-3
: Sequence 3, Application US/08286889
: Patent No. 5470953
: GENERAL INFORMATION:
: APPLICANT: Gallatin, W. Mich
: APPLICANT: Van der Vieren, Monica
: TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive, 6300 Sear Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/286,889
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/173,497
: FILING DATE: 23-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: P38,659
: REFERENCE/DOCKET NUMBER: 27866/32168
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 136
QY 121 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTFLFSLMOYSEEP 180
Db 137 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTFLFSLMOYSEEP 196
QY 181 RHFTFKFQNNPNRSLIKIPITQLGRTHATGLRKVRKVFNITGARKNAFKILFLL 240
Db 197 RHFTFKFQNNPNRSLIKIPITQLGRTHATGLRKVRKVFNITGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGDAFRSEKSRQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTQNLREKIFAIETGTQGTSSSSFEHMSQEGFSAATISNGPILLSSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIETGTQGTSSSSFEHMSQEGFSAATISNGPILLSSTVGSYDWAG 376

361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMPR 420
377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMPR 436
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
481 PRGQARWQCDVAVLGEQGPNGRGAALTVDVGVNGDKLTDAVAGAGEEDNRGAVYLF 540
497 PRGQARWQCDVAVLGEQGPNGRGAALTVDVGVNGDKLTDAVAGAGEEDNRGAVYLF 556
541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600
557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616
601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
677 YDLALDSGPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKKNCNDNICDDLSITFSFMSLCLVVG 780
737 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKKNCNDNICDDLSITFSFMSLCLVVG 796
781 GPREFNVTVVRNDEGDSVRTQVTFPPLDLSYRVKYSTLONQSRQSRWLACASSTEV 840
797 GPREFNVTVVRNDEGDSVRTQVTFPPLDLSYRVKYSTLONQSRQSRWLACASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
901 QLELPVKYAVVWVTVSHGVSTKYNFTASENTSRVWQHGYQVSNLQSRSLPISLVFLVPV 960
917 QLELPVKYAVVWVTVSHGVSTKYNFTASENTSRVWQHGYQVSNLQSRSLPISLVFLVPV 976
961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP 1020
977 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP 1036
1021 PFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABILEFNDVSVFTLLPQGGAFVRSQTEF 1080
1037 PFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABILEFNDVSVFTLLPQGGAFVRSQTEF 1096
1081 KVEPFEPVNPPLPIVGVSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137
1097 KVEPFEPVNPPLPIVGVSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173.497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286.889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362.652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNNSLGLSLAATTSPQLLACGTVHTQTCSENTYVKGCLFGLFNSLNRRQPOK 120
DB 77 RLQVPEAVNNSLGLSLAATTSPQLLACGTVHTQTCSENTYVKGCLFGLFNSLNRRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTFLPSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQLKSKTFLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAPRSEKSRQELMTVASKPPRDRHVFOAN 300
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSRQELMTIASKPPRDRHVFOAN 316
QY 301 NFEALKTVQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVSGYDWMAG 360
DB 317 NFEALKTVQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVSGYDWMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVAVLGEQGPNGRGAALTVDVGVNGDKLTDAVAGAGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLGEQGPNGRGAALTVDVGVNGDKLTDAVAGAGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600

Db 557 HGTSGGSPSHSQR IAGSKLSPLQYFGQSLSGCQLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVWT 660
Db 617 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVWT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLPNCIEDVPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLPNCIEDVPSPVILRLNF 736
Qy 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNSORSQSRWLACASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNSORSQSRWLACASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTBF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTBF 916
Qy 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 977 RLNQTVDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1036
Qy 1021 PFGIQEBFNATLKNLGFWDWIKTSHNHLITVSTAEILFNDSVFTLLPGOGAFVRSQTET 1080
Db 1037 PFGIQEBFNATLKNLGFWDWIKTSHNHLITVSTAEILFNDSVFTLLPGOGAFVRSQTET 1096
Qy 1081 KVEPFEPVNPULIVGSSVGLLLALITAAALYKLGFFKRYKDYKMMSEGGPPGAEPQ 1137
Db 1097 KVEPFEPVNPULIVGSSVGLLLALITAAALYKLGFFKRYKDYKMMSEGGPPGAEPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARFGQS VVQLQGSRVVVGAPQEI VAAHQSGSLYQCDYSTGCEPI 60
Db 17 FNLDTENAMTFQENARFGQS VVQLQGSRVVVGAPQEI VAAHQSGSLYQCDYSTGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGSII PHDFRRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
Db 137 FPALRGCPQEDSDIAFLIDGSGSII PHDFRRMKELVSTIMEOLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTKFEFONNPNRSLKPIITQLLGRTHRTATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 197 RIHFTKFEFONNPNRSLKPIITQLLGRTHRTATGLRKVVRLEFNITNGARKNAFKILV 256
Qy 241 TDCEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFSEKSRQBLNTVASKPPRDHVPOAN 300
Db 257 TDCEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFSEKSRQBLNTVASKPPRDHVPOAN 316
Qy 301 NPEALKTVQNLREKJFAIEGTOTGSSSFEHEMSOEGFSAAITSNGLPSTVSGSDWAG 360
Db 317 NFEALKTIQNLREKJFAIEGTOTGSSSFEHEMSOEGFSAAITSNGLPSTVSGSDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILRNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL 496
Qy 481 PRGORARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAVYLF 540
Db 497 PRGORARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAVYLF 556
Qy 541 HGTSGGSPSHSQR IAGSKLSPLQYFGQSLSGCQLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGGSPSHSQR IAGSKLSPLQYFGQSLSGCQLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVWT 660
Db 617 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVWT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLPNCIEDVPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLPNCIEDVPSPVILRLNF 736
Qy 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNSORSQSRWLACASSTEV 840

Db 797 GPREFNVTVRNDGSDVTRQTVPFFPLDLSVRKUSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTFF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036
Qy 1021 FFGIQEENATLKNLSFDWYIKTSHNHLIIVSTABEILFNDVSFTLLPGGAFVRSQTET 1080
Db 1037 FFGIQEENATLKNLSFDWYIKTSHNHLIIVSTABEILFNDVSFTLLPGGAFVRSQTET 1096
Qy 1081 KVEPPEVNPPLPIVGVSSVGLLLALITAAALYKLGFFKRYKXDMSEGGPPGAEQ 1137
Db 1097 KVEPPEVNPPLPIVGVSSVGLLLALITAAALYKLGFFKRYKXDMSEGGPPGAEQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 127866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-605-672-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAANQORSGLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAANQORSGLYQCDYSTGSCBPI 76
Qy 61 RIQVPEAVNMSLGLSLAATTSPQOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOOK 120
Db 77 RIQVPEAVNMSLGLSLAATTSPQOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOOK 136
Qy 121 FPEARLGCQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKSTLFSLMQYSEBF 180
Db 137 FPEARLGCQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKSTLFSLMQYSEBF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITOLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240
Db 197 RIHFTFKFQNNPNRSLIKPITOLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQAN 316
Qy 301 NFEALKTQVQNLREKIFAIEGTQTCSSSFEHMSQEGFSAAITSNGLPSTVSGSYDWAG 360
Db 317 NFEALKTQVQNLREKIFAIEGTQTCSSSFEHMSQEGFSAAITSNGLPSTVSGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMPR 436
Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYETQTRGQSVVCP 480
Db 437 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYETQTRGQSVVCP 496
Qy 481 PRGQARMQCDALYGEQGPWGRFGAALTVDLVNGDKLTDAIAGAPGEENRGAVYLF 540
Db 497 PRGQARMQCDALYGEQGPWGRFGAALTVDLVNGDKLTDAIAGAPGEENRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRRLRSGQTSVVT 660
Db 617 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRRLRSGQTSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFNGNLRPVLAEDAORLFTALPPFKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGNLRPVLAEDAORLFTALPPFKNGCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGSDVTRQTVPFFPLDLSVRKUSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGSDVTRQTVPFFPLDLSVRKUSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTFF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036

QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFEPNPLPLIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
DB 1097 KVEPFEPNPLPLIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams JK., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPEIQAANQRGSLYQCDYSGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPEIQAANQRGSLYQCDYSGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTKTLFSLWQYSEEF 180

RESULT 7

US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGPTVHTCSENTVYVKGCLFLFGSNLRQQPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGPTVHTCSENTVYVKGCLFLFGSNLRQQPQK 136

Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMQYSEEP 180
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMQYSEEP 196

Qy 181 RIHFTFKFQNNPNRSLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVIPSLEDREGVIRYVLGFDAPFRSEKSRQELNTVASKPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVIPSLEDREGVIRYVLGFDAPFRSEKSRQELNTVASKPRDHVFOAN 316

Qy 301 NFEALKTVQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATSNGLPLLTGVSQYDAG 360
Db 317 NFEALKTVQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATSNGLPLLTGVSQYDAG 376

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 436

Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYIYEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYIYEQTRGGQSVCP 496

Qy 481 PRQARAWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTVDVAIGAPGEDNRGAYVLF 540
Db 497 PRQARAWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTVDVAIGAPGEDNRGAYVLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVLTGAGQHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVLTGAGQHVLRLRSQ 616

Qy 601 PVLRYKAIMFENPREVARNVFECNDQVYVKGKEAGEVRVCLHVQKSTDRDLREGQIQSVVT 660
Db 617 PVLRYKAIMFENPREVARNVFECNDQVYVKGKEAGEVRVCLHVQKSTDRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTCTETLKLQLPNCIEDPVSPIVLRNLF 736

Qy 721 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPFKXKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPFKXKNCNDNICQDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 916

Qy 901 QLELPVKYAVVMVVTSHGVSTKYVLTNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVVTSHGVSTKYVLTNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 976

Qy 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDIP 1020
Db 977 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDIP 1036

Qy 1021 FFGIQEEFNATLKGNTSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGAFVRSOTET 1080
Db 1037 FFGIQEEFNATLKGNTSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGAFVRSOTET 1096

Qy 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match 99.3%; Score 5839; DB 3; Length 1153;									
Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVEAVNMVSLGSLAATTSPPQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	120						
Db	77	RLQVPVEAVNMVSLGSLAATTSPPQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	136						
Qy	121	PPEARLCCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF	180						
Db	137	PPEARLCCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF	196						
Qy	181	RIHFTFEFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITGARKNAFKILFL	240						
Db	197	RIHFTFEFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITGARKNAFKILFL	256						
Qy	241	TGGEKFGDPLGYEDVPEADREGVIRVYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	300						
Db	257	TGGEKFGDPLGYEDVPEADREGVIRVYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	316						
Qy	301	NFEALKTVQNLREKIPAIETGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG	360						
US-09-688-307A-3									
Sequence 3, Application US/09688307A									
Patent No. 6432404									
GENERAL INFORMATION:									
APPLICANT: Gallatin, Michael W.									
TITLE OF INVENTION: No. 6432404el Human Beta-2									
FILE REFERENCE: 27866/36646									
CURRENT APPLICATION NUMBER: US/09/688,307A									
CURRENT FILING DATE: 2000-10-13									
PRIORITY APPLICATION NUMBER: 09/193,043									
PRIORITY FILING DATE: 1998-11-16									
PRIORITY APPLICATION NUMBER: 08/605,672									
PRIORITY FILING DATE: 1996-02-22									
PRIORITY APPLICATION NUMBER: 08/173,497									
PRIORITY FILING DATE: 1993-12-23									
PRIORITY APPLICATION NUMBER: 08/286,889									
PRIORITY FILING DATE: 1994-08-05									
PRIORITY APPLICATION NUMBER: 08/362,652									
PRIORITY FILING DATE: 1994-12-21									
PRIORITY APPLICATION NUMBER: 08/943,363									
PRIORITY FILING DATE: 1997-10-03									
NUMBER OF SEQ ID NOS: 114									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 3									
LENGTH: 1153									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 99.3%; Score 5839; DB 4; Length 1153;									
Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVEAVNMVSLGSLAATTSPPQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	120						
Db	77	RLQVPVEAVNMVSLGSLAATTSPPQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	136						
Qy	121	PPEARLCCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF	180						
Db	137	PPEARLCCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF	196						
Qy	181	RIHFTFEFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITGARKNAFKILFL	240						
Db	197	RIHFTFEFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITGARKNAFKILFL	256						
Qy	241	TGGEKFGDPLGYEDVPEADREGVIRVYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	300						
Db	257	TGGEKFGDPLGYEDVPEADREGVIRVYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	316						
Qy	301	NFEALKTVQNLREKIPAIETGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG	360						

Db 317 NFEALKTIONQUREKIPAIETGQTGSSSFHEHMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVCP 496
Qy 481 PRGQARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 616
Qy 601 PVLVRKAIIMEFNPREVARNVFECNDQVWKGKAGEVRVCLHVQKSTRDLRGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECNDQVWKGKAGEVRVCLHVQKSTRDLRGQIQSVVT 676
Qy 661 YDLALDSGPHRAVENETKNSRTRQTVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGPHRAVENETKNSRTRQTVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKNCNDNICODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKNCNDNICODDLSITFSFMSLCLVVG 796
Qy 781 GPREFNVTVTRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVTRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCINHIPIPPENSEVTNITFPDVSASLGKGLKLLKANVTSENWPRNKTET 900
Db 857 SGALKSTSCINHIPIPPENSEVTNITFPDVSASLGKGLKLLKANVTSENWPRNKTET 916
Qy 901 QLELPVKYAVVMVTSHGVSSTKYLNTASNTSRVMQHOYQVSNLQORSPLISLFLVLPV 960
Db 917 QLELPVKYAVVMVTSHGVSSTKYLNTASNTSRVMQHOYQVSNLQORSPLISLFLVLPV 976
Qy 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQC 1020
Db 977 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQC 1036
Qy 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQCGAFVRSQ 1080
Db 1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQCGAFVRSQ 1096
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKROKDMWSEGGPPGABPO 1137
Db 1097 KVEPPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKROKDMWSEGGPPGABPO 1153

RESULT 10

US-08-476-062A-43

Sequence 43, Application US/08476062A

Patent No. 5877275

GENERAL INFORMATION:

APPLICANT: Arnaout, M. Amin

TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAQNRSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARFGQSVVQLQGSRVVVGAPQEIIVAAQNRSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVVKGLCFGLFNLNQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVVKGLCFGLFNLNQPOK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLI 256
Qy 241 TDGEKFGDPLGYEDVIPEDLREGVIRYVYVGLFGDAPFRSEKSRQELNTVASKPPDRHVFOAN 300
Db 257 TDGEKFGDPLGYEDVIPEDLREGVIRYVYVGLFGDAPFRSEKSRQELNTVASKPPDRHVFOAN 316
Qy 301 NPEALKTVONOLREKIPIAIEGTQGTSSSFHEHMSQEGFSAAITNSGPLLSTVGSYDMAG 360
Db 317 NPEALKTVONOLREKIPIAIEGTQGTSSSFHEHMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQSVCP 496
Qy 481 PRGQARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600

Db 556 HGTSGGSPSHSQRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVAAGHVLRLRSQ 615
Qy 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGIOISVWT 660
Db 616 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGIOISVWT 675
Qy 661 YDLALDSGRPHSRVAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDVPSPVLRLNF 720
Db 676 YDLALDSGRPHSRVAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDVPSPVLRLNF 735
Qy 721 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCVLVVG 780
Db 736 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCVLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 915
Qy 901 QLELPVKYAVTVMTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRPLISLVFLVPV 960
Db 916 QLELPVKYAVTVMTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRPLISLVFLVPV 975
Qy 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020
Db 976 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1035
Qy 1021 FFGIOEENATLKNLSFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1080
Db 1036 FFGIOEENATLKNLSFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1095
Qy 1081 KVEFEVNPPLIVGSSVGGLLIALITALYALYKLGFFKQYKDMSEGGPPGAEPQ 1137
Db 1096 KVEFEVNPPLIVGSSVGGLLIALITALYALYKLGFFKQYKDMSEGGPPGAEPQ 1152

RESULT 11

PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906

TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 98.8%; Score 5808.5; DB 5; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNRQSLYQCDYSTGSCPEI 60

Db 17 FNLDTENAMTFFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNRQSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVYKGLCFLFGSNLRQPOK 120

Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVYKGLCFLFGSNLRQPOK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFRRMKELVSTIMEOLKSKTKLFSLMQYSEEF 180

Db 137 FPEALRGCPQEDSDIAFLIDGSGIIPHDFRRMKELVSTIMEOLKSKTKLFSLMQYSEEF 196

Qy 181 RIHFTFEKFNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLL 240

Db 197 RIHFTFEKFNPNPRSLIKPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAFRSEKSRQELNVTASKPPRDHVFQAN 300

Db 257 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAFRSEKSRQELNVTASKPPRDHVFQAN 316

Qy 301 NFEALKTQNLREKIPAEIGTOTGSSSPSEHEMSQEGFSAATTSNGPLLSTVGSYDWAQ 360

Db 317 NFEALKTQNLREKIPAEIGTOTGSSSPSEHEMSQEGFSAATTSNGPLLSTVGSYDWAQ 376

Qy 361 GVFLYTSKEKSTFINMTRVDSNMNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMFR 420

Db 377 GVFLYTSKEKSTFINMTRVDSNMNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPPL 480

Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPPL 496

Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAVYLF 540

Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAVYLF 555

Qy 541 HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVAAGHVLRLRSQ 600

Db 556 HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVAAGHVLRLRSQ 615

Qy 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGIOISVWT 660

Db 616 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGIOISVWT 675

Qy 661 YDLALDSGRPHSRVAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDVPSPVLRLNF 720

Db 676 YDLALDSGRPHSRVAVFNETKNSRROTQVLGLTCTETLKLQLPNCIEDVPSPVLRLNF 735

Qy 721 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCVLVVG 780

Db 736 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCVLVVG 795

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

Db 796 GPRESNVTVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 855

Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900

Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 915

QY 901 QLELPVKYAVVMTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVVPV 960
DB 916 QLELPVKYAVVMTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
DB 976 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1035
QY 1021 PFGIOEENATLKGNSLSPDWMYIKTSHNHLIIIVSTAEILFNDSPVFTLLPQGGAFVRSQTET 1080
DB 1036 PFGIOEENATLKGNSLSPDWMYIKTSHNHLIIIVSTAEILFNDSPVFTLLPQGGAFVRSQTET 1095
QY 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
DB 1096 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2:
; LENGTH: 1152
5424399-2

Query Match 98.8%; Score 5808.5; DB 6; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVLOGSRVVVVGAPQEIIVAAHQRSGLYOCYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVLOGSRVVVVGAPQEIIVAAHQRSGLYOCYSTGSCBPI 76
QY 61 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPOK 120
DB 77 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSBEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSBEP 196
QY 181 RTHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEXFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNNTVASKPPRDRHVQAN 300
DB 257 TDGEXFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNNTIASKPPRDRHVQAN 316
QY 301 NFEALKTVQNRKIFAJEQTQSGSSFEHEMGOEFGSAITNGPILLSVTGSDVWAG 360
DB 317 NFEALKTIQNLKRIKIFAJEQTQSGSSFEHEMGOEFGSAITNGPILLSVTGSDVWAG 376
QY 361 GVFLYTSKESKTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYOHIGLVAMFR 420
DB 377 GVFLYTSKESKTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 496
QY 481 PRQBARWQCDVLYGEOQPMWRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

DB 497 PRG-RARWQCDVLYGEOQPMWRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDTLVGAQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDTLVGAQGHVLLRSQ 615
QY 601 PVLRYKATMEPNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRRLREGOIQSVVT 660
DB 616 PVLRYKATMEPNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRRLREGOIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNF 720
DB 676 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNF 735
QY 721 SLVGTPLSAFGNLRPLVLAEDAQRLLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFGNLRPLVLAEDAQRLLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVVRNDEGDSYRQVTFPFLDLSYRKVSTLQORSQSRWLACESASSTEV 840
DB 796 GPRESNVTVVRNDEGDSYRQVTFPFLDLSYRKVSTLQORSQSRWLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVVPV 960
DB 916 QLELPVKYAVVMTSHGVSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
DB 976 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1035
QY 1021 PFGIOEENATLKGNSLSPDWMYIKTSHNHLIIIVSTAEILFNDSPVFTLLPQGGAFVRSQTET 1080
DB 1036 PFGIOEENATLKGNSLSPDWMYIKTSHNHLIIIVSTAEILFNDSPVFTLLPQGGAFVRSQTET 1095
QY 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
DB 1096 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1152

RESULT 13
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

Qy	1	F	N	L	D	T	E	N	A	M	T	F	Q	E	N	A	R	G	F	G	S	V	Q	L	O	G	S	R	V	V	V	G	A	P	O	I	V	A	A	N	O	R	G	S	I	Y	Q	C	D	I	S	T	G	S	C	E	P	I	60
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	79				
Qy	61	R	L	O	V	P	E	A	V	N	S	L	G	L	S	I	A	A	T	T	S	P	O	L	L	A	C	P	T	V	H	O	T	S	E	N	T	Y	K	G	L	C	F	L	G	S	N	L	R	O	O	P	K	120					
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	137				
Qy	121	F	P	E	A	L	R	C	O	P	E	D	S	I	A	I	L	D	G	S	G	I	I	P	H	D	F	R	R	K	E	L	V	T	I	M	E	O	L	K	K	S	T	L	F	S	L	M	O	Y	S	E	E	F	180				
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	197				
Qy	181	R	I	I	H	T	F	K	E	P	O	N	N	P	R	S	L	I	K	P	I	T	O	L	L	O	R	T	H	A	T	G	L	R	K	V	R	E	L	F	N	I	T	N	G	A	R	K	N	A	F	K	L	F	L	240			
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	257				
Qy	241	T	D	E	K	F	G	D	P	L	G	E	D	V	I	P	E	L	O	R	E	G	V	I	R	V	V	L	G	F	D	A	R	S	E	K	S	R	O	E	L	N	T	V	A	S	K	P	P	R	D	H	V	F	O	A	N	300	
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	317				
Qy	301	N	F	E	A	L	T	V	Q	N	L	R	E	K	I	F	A	I	E	G	T	O	G	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	360	
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	377				
Qy	361	G	V	F	L	T	S	K	E	K	S	T	F	N	M	T	R	V	D	S	D	M	N	D	A	Y	L	G	A	A	I	I	L	N	R	V	O	S	I	V	L	G	A	P	R	Y	O	H	I	G	L	V	A	M	F	R	420		
Db			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	437				
Qy	421	Q	N	T	G	M	E	S	N	A	N	V	G	T	O	I	G	A	F	G	A	S	L	C	S	V	D	S	N	G	S	T	D	L	V	L	I	G	A	P	H	Y	E	O	T	R	G	O	V	S	C	P	L	480					
Db			:	:	:	:	:																																																				

[illegible]

Db 318 DFDALDKIQNLKEKIFAIECTETISSSSFELEMAQEGSAVFTPDGPVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLYGAAAAIILNRVQSLVGLGAPRYQHICGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVGLGAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGQGVSVCP 480
Db 438 QVSRQWENKAEVITQIGSYFGASLCSVDVDTGSTDVLIGAPHYYEQTGQGVSVCP 497
Qy 481 PRGORARQCDAVLYGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYL 540
Db 498 PRGWR-RWMCDAVLYGQGPWGRFGAALTVLGVDVNGDKLTDVVGAPGEEENRGAYLF 556
Qy 541 HGTSGSGLSPSHSRIAGSLSPLOYFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPHSRIAGSLSSRLQYFGQALSGQDITQDGLVDLAVGARGQVLLRTR 616
Qy 601 PVLRVKAIMENPREVARNVECNQVVKGEAGEVRVCLHVOKSTRDRLREGQIQSWT 660
Db 617 PVLWGVSMQIPAEIPRSAFECEQVVEQTLVQSNICLYIDKRSKNLGSRLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRRTQVGLGTQTCETLKLQLPNCIEDPVSPVLRLNF 720
Db 677 LDLALAPGLSPRAIFQETKNSRVRVGLKAKHCENFNLLPSCVEDSVIPIILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLFTALFPFEKNCNDNICODDLISITFSFMSLDCLVWG 780
Db 737 TLVGKPLAFNLRLPMLAALAQRYFTASLPFEKNCAGADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPREFNVTVVRNDEDSYRTQVTFPPDLDSYRKVSTLQNRORSORWRLACESASSTEV 840
Db 797 SNLELNAEVMWNGEDSYGTTITFHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 855 SGTWSTSCRINHILFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVYVWVTSKYNLFTAS-ENTSRVMOHQYVNSLQORSPLISLVLP 959
Db 915 QLELPVKYAVYVWVSSHQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQIOCDI 1019
Db 975 VELNQEAVWMDVEVSHPONPSLRCSSEKIAPPASDFLAHIOKNPVLDCSIAGCLRFRCDV 1034
Qy 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1079
Db 1035 PSFSVQEELDFTLKGNLSFGWVROILQKKSVSVVAEIIFDTSVYSQLPQGEAFMRAQTI 1094
Qy 1080 TKVEPFEVNPPLIVGSSVGGLLLALITALYKLGFFKQYKDMHSE 1128
Db 1095 TVLEKYVHNPIPLIVGSSIGGLLLALITALYKVGFFKQYKEMMEE 1143

Search completed: November 25, 2003, 14:23:13
Job time : 17.5618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFOENARGFCQ.....FKROYKDMSEGPGCAEPO 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 RWHULB	cell surface glyco
2	4460	75.8	1133	2 S00551	leukocyte surface
3	3469	59.0	1153	1 RWHULC	cell surface glyco
4	1543.5	26.2	1170	2 S03308	cell surface glyco
5	1527.5	26.0	1163	2 I56126	lymphocyte fuction
6	1140	19.4	1179	2 A53213	integrin alpha-E c
7	1093.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45314	integrin alpha 2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1055	17.9	1180	2 A35854	integrin alpha-1 c
11	1054	17.9	1181	2 A33998	integrin alpha-2 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.4	1041	2 T31437	integrin alpha-1 c
16	579.5	9.8	1054	2 JC7294	alphanp integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.6	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1033	2 S44250	integrin alpha-5 c
20	546.5	9.3	1034	2 A36108	integrin alpha-v c
21	535	9.1	1044	2 T10050	integrin alpha-6 c
22	534.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	532	9.0	1073	2 B36429	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	8.9	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

RESULT 1

RWHULB

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C;Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; I52567

R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,
B.

A;Reference number: A31108; MUID:88315033; PMID:2457584

A;Accession: A31108

A;Molecule type: mRNA

A;Residues: 1-1153 <COR>

A;Cross-references: GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:G307148

A;Note: part of this sequence was confirmed by protein sequencing

R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor:
A;Reference number: A28915; MUID:88257215; PMID:2454931

A;Accession: A28915

A;Molecule type: mRNA

A;Residues: 1-499,501-965 'P', 967-1153 <ARN>

A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA-
A;Note: The authors translated the codon TAC for residue 1129 as Thr

A;Note: part of this sequence, including the amino end of the mature protein, was conf
R;Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally
A;Reference number: A41600; MUID:92073318; PMID:1683702

A;Accession: A41600

A;Molecule type: DNA

A;Residues: 1-9 <SHE>

A;Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:G553215

R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
A;Reference number: A94193; MUID:88190151; PMID:2833753

A;Accession: A30892

A;Molecule type: mRNA

A;Residues: 917-1042 <AR2>

A;Cross-references: GB:M18044

R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept
A;Reference number: A32218; MUID:89098893; PMID:2563162

A;Accession: A32218

A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>

A;Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:G386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Molecule type: DNA
A:Status: not compared with conceptual translation
A:Residues: 1-499, 501-1153 <FL>
A:Cross-references: GB:S52227; NID:9263047; PIDN:AA24821.1; PID:9263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26031
A:Molecule type: protein
A:Residues: 17-31 <PI>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:9553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,977,735,808,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVVOLQSGSRVVGAPQEIIVANQSGSLYQCYDSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVOLQSGSRVVGAPQEIIVANQSGSLYQCYDSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVGCLFLGSLNRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVGCLFLGSLNRQPOK 136

Qy 121 FPEALRGCPQSDSIAFLIDGSGSIIIPHDPRMKEWSTVMQELKSKTFLSLMYSEEF 180
Db 137 FPEALRGCPQSDSIAFLIDGSGSIIIPHDPRMKEFVSTVMQELKSKTFLSLMYSEEF 196

Qy 181 RIHTFFKFNPNPNSRIKPIITQLLGRTHATGLRKVRELFINITNGARKNAFKILFL 240
Db 197 RIHTFFKFNPNPNSRIKPIITQLLGRTHATGLRKVRELFINITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGVEDVPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPDRHVFOIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSKQELNTVASKPPDRHVFOVN 316

Qy 301 NFEALXTIQNLREKIFAIEGTOTGSSSSFEHEMSQEGFSAALTSNGPLLLSTVGSYDWAG 360
Db 317 NFEALXTIQNLREKIFAIEGTOTGSSSSFEHEMSQEGFSAALTSNGPLLLSTVGSYDWAG 376

Qy 361 GVFLYTSKSKSTFINNTRVDSQNDMDAYLGVAALILNRNVQSLVLCAPRYQHILGLVAMPR 420
Db 377 GVFLYTSKSKSTFINNTRVDSQNDMDAYLGVAALILNRNVQSLVLCAPRYQHILGLVAMPR 436

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGRGQGVSCPL 496

Qy 481 PRQQRARWQCDVLYGECQCPWGRFGAALTVDGVNGDKLTDVAICAPGEEDNRGAVYLF 540
Db 497 PRQQRARWQCDVLYGECQCPWGRFGAALTVDGVNGDKLTDVAICAPGEEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSIVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSIVLRNLF 736

Qy 721 SLVGTPLSAFGNLRPVLAEADAQRLFPPEKKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEADAQRLFPPEKKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVTRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSQRSLACSSASSTEV 840
Db 797 GPREFNVTVTRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSQRSLACSSASSTEV 856

Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVTSENNMPTNKTEF 916

Qy 901 QLSELPVKYAVYVMTSHGVSTKYNLTASNTSRVMOHQVQVSNLQORSPLISLVFLVPV 960
Db 917 QLSELPVKYAVYVMTSHGVSTKYNLTASNTSRVMOHQVQVSNLQORSPLISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1036

Qy 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGGAFVRSQTE 1080
Db 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGGAFVRSQTE 1096

Qy 1081 KVEPFEPVNPPLIIVGSSVGLLLALLITAALYKLGFFKRYKQKXDMMSGEGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLIIVGSSVGLLLALLITAALYKLGFFKRYKQKXDMMSGEGPPGAEPQ 1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N:Alternate names: complement-3 receptor alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C:Accession: S00551; I59078
R:Pytela, R.
EMBO J. 7, 1371-1378, 1988
A:Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th
A:Reference number: S00551; MUID:88312584; PMID:3044779
A:Accession: S00551
A:Molecule type: DNA
A:Residues: 1-1153 <PYT>
A:Cross-references: EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:952983
A>Note: the authors translated the codon CAC for residue 569 as Gln

R.Saetre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Cross-references: GB:M14293; NID:q198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;
Best Local Similarity 73.8%; Pred. No. 2.2e-303;
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

Qy 1 FNLDTENAMTFQENARGFQSVVQJGSRVTVVGAQOEIVAAHQRSGLYQCDYSTGSCBPI 60
Db 17 FNLDTEHPMTFQENAKFGQNVVQLGGTSVVVAAPQEAQAVNQTGALYQCDYSTGRCHI 76

Qy 61 RLQVPVEAVNMSIGLSLAATTPPOLACGPVTHQTCSENTYVKGICFLFGSNLRQQPOK 120
Db 77 PLQVPPEAVNMSIGLSLAATTPPOLACGPVTHQNCENTYVNGLCYLFGSNLRPPQO 136

Qy 121 FPEALRGCPQEDSDTAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMQYSBEF 180
Db 137 FPEALRECPQESDIVFLIDGSGSINNIDFQMKFEFVSTVMEQFKSKTFLSLMQYSDEF 196

Qy 181 RIHFTKFEQNNPNSRLIKPTQLGLRTHATGKRVVRELNFNITNGARKNAFKILFL 240
Db 197 RIHFTFNDPKRPSRSHVSPKQLNGRTKTASGIRKVVRELPHKNTGARENAKILVVI 256

Qy 241 TQGEKFGDPLGYEDVPELDRGVIRYVIGDGAFRSEKSRQELNATVASKPRDHVFQIN 300
Db 257 TDGEXFGDPLDYKVIDPEADRAGVIRYVIGVGNANFNKQSRRELDTITASKPAGEHVQVD 316

Qy 301 NFEALKTIQNLQREKIFALEGTQTSSEHSEMSQEGFSAITNGPLLSVTGVSVDWAG 360
Db 317 NFEALNTIQNLQEKIFALEGTQTSSEHSEMSQEGFSAITNGPLLSVGSVDFWAG 376

Qy 361 GVFLYTSKESKTFINNRVDSNDNDAYLGYAAAILNRNVQSLVIGAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTNRVDSNDNDAYLGYASAVILNRNVQSLVIGAPRYQHIGLVAMFR 436

Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAPHYETRGQGVSCPL 480
Db 437 ENFGTWEPHTSTKSGQISGYFGASLCSVDMDADGNTNLILGAPHYETRGQGVSCPL 496

Qy 481 PRGQARMQCDVAVLGEQGPNGRFGAALTVLGVNGDKLTDVAIGAPDEENRGAVILF 540
Db 497 PRG-RARWQCEALLHGDQHPGRFGAALTVLGVNGDKLTDVAIGAPDEENRGAVILF 555

Qy 541 HTSGSGISPSHSQRTAGSKLSPRLQYFQCSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 556 YGASIASLSSASHRIIGAHFSPGLQYFQCSLGGQDLTMDGLMDLAVGAQGHVLLRSQ 615

Qy 601 PVLRVKAIHFNPREVARNVFECNDQVVGKAGEVRVCLHVOKSTRDLRGQTSVVT 660
Db 616 PVLRLLEATMEFSPKPKVARSVFACQEQVLKNKDAGEVRVCLRVKNTKDLRGDQTSVVT 675

Qy 661 YDLALDGRPHARVFNETKSTRQTVLGLTQTCETIKLQLPNCIEDPVPVILRLNF 720
Db 676 YDLALDVPKSRIRAFDETKNTRRTQVFLGQKQKCTLKLTLPOCDVDSVPFIIILRLNY 735

Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPFPFKNCGNDNICDDLSITFSPMSLCLVVG 780
Db 736 TLVGEPLRSFGLNLRPVLANDQARFFFTAFPFKFCNCGNDSICQDDLSITMSANGLDVLVG 795

Qy 781 GPREFNVTVTVRNDEGDSYRTQVTFPPDLDSYKRVKSTLQNRQSRQSRML-ACESASSTE 839
Db 796 GPQDFNMSVTLNRNDEGDSYGTQVTVYYPGSLSYKSDASQNPITKPKPWFVPAESSSSE 855

Qy 840 VSGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKLLKANVTSENMRPTNTE 899
Db 856 GHGALKSTTWNINHPIIPANSEVFNITFDVDSHASFGNKLKLLKAIIVASENNMSRTHKTK 915

Qy 900 FQLELPVKYAVYVMTVSHGVTSTKYLNFNTASENTSRVMOHOYQVSNLQORSPLISLVLVP 959
Db 916 FQLELPVKYAIYMTVTSDESSIRYLNFTASMTSKVIOHQYQFNNGQORSUPVSVFWFIP 975

Qy 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDI 1019
Db 976 VQINNVTWMDHPQVIFSQNLSSACHTEQKSPHNSFRDOLERTVPLNCSVAVCKRIQCDL 1035

Qy 1020 PFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db 1036 PSPNTQEIFNVTLKNLSFDWYIKTSHGHLVLSSTELFNDSAPALLPGQESVYRSKTE 1095

Qy 1080 TKVSPFEVNPPLIVGVSSVGLLLALITAAALKYKGFKKQYKDMMSGGPPGAEPQ 1137
Db 1096 TKVSPYEVNVPPLIVGVSSIGLVLLALITAGLYKGFKKQYKDMMEAAAPQDAPPQ 1153

RESULT 3
RWHUIC
cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C:Accession: A36584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <OR>
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:86166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:q487829; PIDN:AAA59180.1; PID:q487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat ho
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repe
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61.89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match		59.0%	Score 3469;	DB 1;	Length 1163;
Best Local Similarity		60.9%	Pred. No. 5e-234;		
Matches 688;		Conservative 142;	Mismatches 293;	Indels 6;	Gaps 4;
Qy	1	FNLDTENAMTFQENARFGQSVVOLQGSRRVVGAPQEIIVAAVNRQSGLSYQCDYDSTGSCPEI	60		
Db	20	FNLDTENAMTFQENARFGQSVVOLQGSRRVVGAPQEIIVAAVNRQSGLSYQCDYDSTGSCPEI	79		
Qy	61	RLQVPVAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK	120		
Db	80	GLQVPPVAVNMSGLSLAATTSPPOLLACGPTVHQBGRNMYLGLCFLGLGPT--QUTQR	137		
Qy	121	FPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRRKKEWSTVMEQKKSKTLFSLMQSSEEF	180		
Db	138	LPVSRQECPCQEQDIFVLIDGSGSISRNFAFMNFVRAVISQFQRPSTQFSLMQSFKF	197		
Qy	181	RIHFTPEFQNNPNRSLIPITOLLGRTHTATGLRKVRELFNITNGARKNAFKILFL	240		
Db	198	QTHFTFEFRRTSNPLSLLASVHOLQFTYTATAIQNVVHRLFHASYGARADATKILIVI	257		
Qy	241	TDGEKFGDPLGYEDVPELDRGVIRYVIGVDGDAFRSEKSRQELNVTASKPPRDHVOIN	300		
Db	258	TGKKEGSDLDYKDVIPMAAAGIIRYVIGVLAFOQRNKNWKEINDIASPSQEHIFKVE	317		
Qy	301	NFEAKTIQNLREKIFAIEGTQTSSTSSSEHEMSQCFSAAITNSGPELLSTVGSYDWAG	360		
Db	318	DFDALKDIQNLREKIFAIEGTETTSSTSSSEHEMSQCFSAITPDGVLGAVGSGFTMSG	377		
Qy	361	GVFLYTSKSKSTFINMTVSDSDMDNAYLGYAAAIILNRVOSLVGLAPRYOHGLVAMFR	420		
Db	378	GAFLYPPNMSPTFINMSQENVDMDSDYLGYSTELALWKVQSVLGLAPRYOHTGKAVIFT	437		
Qy	421	QNTGMWESNANVKCTOIGAVFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPL	480		
Db	438	QVSRQWMAKEVCTOIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQGVSCPL	497		
Qy	481	PRGORARWQCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYL	540		
Db	498	PRGWR--RWCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVVGAPGEENRGAYL	556		
Qy	541	HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVLRLRSQ	600		
Db	557	HGVLGPIISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLAVGARGVLLLRTR	616		
Qy	601	PVLAVKAIMENPREVARNVPECNDQVVKGEAGEVAVCVLHVOKSTRDRREGIOISVVT	660		
Db	617	PVLAVGVMQFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLGSRLQGSVT	676		
Qy	661	YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPITVLRNF	720		
Db	677	LDLALDFGLSPRATFOETKRSLSRVRLGLKAHCENFNLLLPSCVEDSVTPITLRLNF	736		
Qy	721	SLVGTPLSAFNLAPVLAEDQRULFTALFPPEKKNCGNDNICQDDLSITFSFMSLDCLVWG	780		
Db	737	TLVGKPLAFNLAPVLAEDQRULFTALFPPEKKNCGADHICQDNLGIFSFPGLKSLVVG	796		
Qy	781	GPRFNVTVVRNDEGDSYRQVTFPFPPLDLSYKSVSTLQNRQSORSWRLACESASSTEV	840		
Db	797	SNLENAEVMWMDGSDSYGTTITFSPAGUSYRYVAEGQKQGLRSLHLTCDSPVGV--	854		
Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900		
Db	855	SQGTWSTSCRINHPIFGGAQITFLATFDVSPKAVLGDRLULLTANVSENNTPTSKTF	914		
Qy	901	QLELPVKYAVYVMTSGVSTKYLNFTAS--ENTSRVMQHOYQVSNLQORSPLISLFLVP	959		
Db	915	QLELPVKYAVYVMTSGVSTKYLNFTAS--ENTSRVMQHOYQVSNLQORSPLISLFLVP	974		
Qy	960	VLNQTWIDRPOVTFPENSLSSTCHTKERLPSHSDFLAELRKAPVNVCSITAVCORICDI	1019		
Db	975	VELNOEAVMMDVESHQNPNSLRSSEKXIAPPASDFLAHQKNVPLDCSIAGCULRFCDV	1034		
Qy	1020	PPFGTQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE	1079		

Db	1035	PSFSVQEEEDFTLTKGNLSPGWRQILQKKVSVVVAEITFDTSVVSQLPQCAFMRQTT	1094		
Qy	1080	TKVEPEVPEPNPLIVGSSVGGLLALALITALYALYKLGFFKROYKOWMSE	1128		
Db	1095	TVLEKYKHVNPTPLIVGSSIGSGLLLALITALYLVYKVGFFKROYKOWMSE	1143		
RESULT 4					
S03308					
cell surface glycoprotein CD11a precursor - human					
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function					
C;Species: Homo sapiens (man)					
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 20-Aug-1999					
A;Accession: S03308; A47565; A48759; S36044					
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.					
J. Cell Biol. 108, 703-712, 1989					
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subu					
A;Reference number: S03308; MUID:89139587; PMID:2537322					
A;Accession: S03308					
A;Molecule type: mRNA					
A;Residues: 1-1170 <LAB>					
A;Cross-references: EMBL:Y00796; NID:g31421; PIDN:CAA68747.1; PID:g31422					
A;Note: Part of this sequence was confirmed by protein sequencing					
R;Cornwell, R.D.; Gollahan, K.A.; Hickstein, D.D.					
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993					
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) 1					
A;Reference number: A47458; MUID:93248261; PMID:8097887					
A;Accession: A47458					
A;Molecule type: DNA					
A;Residues: 1-20 <COR>					
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)					
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.					
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993					
A;Title: Identification of cell-specific and developmentally regulated nuclear factor:					
A;Reference number: A47565; MUID:93281759; PMID:8099450					
A;Accession: A47565					
A;Molecule type: DNA					
A;Residues: 1-20 <SHE>					
A;Cross-references: GB:M95609					
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.					
J. Biol. Chem. 268, 19305-19311, 1993					
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.					
A;Reference number: A48759; MUID:93374910; PMID:8103515					
A;Accession: A48759					
A;Molecule type: DNA					
A;Residues: 1-20 <NUE>					
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406					
C;Genetics:					
A;Gene: GDB:ITGAL; CD11A					
A;Cross-references: GDB:119757; OMIM:153370					
A;Map position: 16p11.2-16p11.2					
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h					
F;1-25/Domain: signal sequence #status predicted <SIG>					
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted					
F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>					
Query Match 26.2%; Score 1543.5; DB 2; Length 1170;					
Best Local Similarity 34.3%; Pred. No. 2.8e-99;					
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;					
Qy	1	FNLDTENAMTFQ--ENARGFGQSVVOLQGSRRVVGAPQEIIVAAVNRQSGLSYQCDYDSTGSCSE	58		
Db	26	YNLDVVGARFSPPRAGRHFYGVLVQV--GNGVIVGAPGE---GNSGTSLYQCQSGTGHCL	81		
Qy	59	PRLQVPVAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLR---	115		
Db	82	PVTLR--GSNYSYKVLGWTLATDPTDGSILLACDPLSRTCDQNTYLSGLCYLFRQNLQPM	140		
Qy	116	-QOPQKPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRRKKEWSTVMEQKKSKTLFSLM	174		
Db	141	LQRPQGFQECIKG---NVDLVFLFDGMSLQPDDEFQKILDFMDKMKKLSNTSYQFAAV	196		

A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBI:124326)
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F:142-317/Domain: von Willebrand factor type A repeat homology <VWAL>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 8.4e-68; Indels 203; Gaps 44;
Matches 345; Conservative 206; Mismatches 487;

QY 1 FNLDTENAMTFOENARG-FCQSVVLQ---QGSRVVVGAPOEIVAAQNRGLYQCDYSTGS 56
DB 1 FNVVDKNSMTFSGPVEDMEGYTVQQYENEEGKWLLIGSLVGPQKNRTGDVYKCPVGRGE 60

QY 57 CEP-IRLQVPEA-----VNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGQFLACGPLYAYRCGHLHYTTGI 119

QY 107 CFLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKK 166
DB 120 CSDVPTFQVNSIAP-VQECSTQ-LDIVLDGNSIYPWDS-----VTAFLNDLLK 170

QY 167 -----SKTLFSLMOYSEEFRIHFTPKFQNNPNRSLIKPITQLLGR-THTATGLRKVV 219
DB 171 RMDIGPKQTVGIVQYGENVTHEFNLNKYSSTEEVLVAAKKIVQGRQTMALGTDITAR 230

QY 220 RELFNITNGAKRNAPKILFLDIDGKEGPGPLGVEDVIPELDREGVIRYVIGVDAPR--- 276
DB 231 KEAFTEARGARRGKVKVMVIVTDGESH-DNHRLLKKVIQDCEDENIQRFSAIILGYSNRGN 289

QY 277 --SEKSRQELNTVASKPRDRHQFNQINNFEALKTIONLREKIPATEGTQSGSSSPEHEM 334
DB 290 LSTEFVEIKSIASEPTEKHFNFVSDLEALVTIVKTLGERIFALEATADQAAAFEMEM 349

QY 335 SOEGFSAATISNGPLLSLVGSDWAGGVPLYTSKE-----KSTF-INMTRVDSMDNDAYL 388
DB 350 SQTGFSAHYSQDWMLGAVGDMNGTVVMQXASQIIIPRNTTFNVESKNEPL-ASVL 408

QY 389 GYAAAILNRRVQSL-VLGAPYQHIGLVAMPQNTGMWESANVKGTOIGAYFASLCS 447
DB 409 GYTVNSATASSGDVLYIAGQPRYNTGQVIIRMEDGNKIKQLTSLGEGIGSYFSGILTT 468

QY 448 VVDVDSNGSDVLVIGAPHY-----YEQTR-GGOVSVCPPLQRCQARWQCD 491
DB 469 TDIDKDSNTDILLVGAPMTYMGTEKEQGVYVVALNQTRFEYQMSLEPIKQCCSSRQHN 528

QY 492 AVLYGEOQOPWG-RFGAALTVLDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISP 550
DB 529 SCTTENKNEPCGARFGTALAAVKDLNLDGFNDIVIGAPLEDHGGAVYIYHG-SGKTIRK 587

QY 551 SHSQRTAGSKLSPRLQYFCQSLSGGODLTWMDGLVDLTVAQGHVLLRSLQPVLRVKAIN 610
DB 588 EYAQRIPSGDGLTKLFFQSGIHGEMDNGDGLTDVTIGLGAALFWRSDVAVVKTWN 647

QY 611 FNPREVARNVFECDNVVKGKAG--EVRVCLHVQ-KSTRDLRLEQIQSVVYDIALDS 667
DB 648 FEFKNVIQKNCH--MEGETVCINATVCFEVLKSKEDTIYEADLQ-----YRVLDS 700

QY 668 GRPHSRVFNET-----KNSTRQTOVLGLTQTCETLKLQLPNCIEDPVSPVLRLNFS 721
DB 701 LQQISRSPSGTQERKQVQNRITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFN 752

QY 722 LVGTPLUSAFGNLRPVLAEDAQRLFTALFPFEKNCNDNTICODLSITTFMSLDCLVWGG 781
DB 753 LT-DPENG-----PVLDSDSPNSVHEYIPFAKDCGNKKECISDLSLHVATTEKDLIVRS 806

QY 782 PRE-FNVTVVRNDGSDSRVQTQVTFPPPLDLSYRKVSTIQNQRSQBSWLACESASTVE 840
DB 807 QNDKFNVSITVKNTKDSAYNTRITVHYSPLNLFSGIEATQKO-----SCSEN----- 853

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (catle)

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C:Accession: I45914

R:Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integr

A:Reference number: A54402; MUID: 94193647; PMID: 7511592

A:Accession: I45914

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <KAM>

A:Cross-references: GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g439696

C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F:161-336/Domain: von Willebrand factor type A repeat homology <VWAZ>

Query Match

Best Local Similarity 18.2%; Score 1071; DB 2; Length 1170;

Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFO-ENARGFGQSVVLQ---QGSRVVVGAPOEIVAAQNRGLYQCD--DYST 54
DB 19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSFPGPKNRMGDVYKCPVDLST 78

QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKG 107
DB 79 TTCEKLNLOTSTSMNVTMTNMSLGLTLTRNVTGGLTCGFLWAOQCGSOYTTGVC 138

QY 108 FLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLK-- 165
DB 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVWCDSENSIYPWD--AVKNFLEKPVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTPKFQNNPNRSLIKPITQLL----GRTHATGLRKVRE 221
DB 195 PTKTQWGLIQIYANNPRVFNLTNTPKSKD---EMIKATSQTFOYGGDLTNTFKAIQYARDT 251

QY 222 LENITNGARKAFKILFLLTDGKEGDPGLVEDVIPELDREGVIRYVIGV-----GDAPR 276
DB 252 AYSTAAGRPAGTKVMVYVVTGESH-DGSKLKAVIDQCNKNILRFGIYAVGLYLRNALD 310

QY 277 SEKSRQELNTVASKPRDRHQFNQINNFEALKTIONLREKIPATEGTQSGSSSPEHEMSQ 336
DB 311 TKNLKKEIKAIASIPTEHFFNVSDADLLEKAGTIGEOIFSIETVQCG-GNPFQNEMSQ 369

QY 337 EGFSAAIT--SNGPLLSLVGSDWAGGVPLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391
DB 337 EGFSAAIT--SNGPLLSLVGSDWAGGVPLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391

370 VGFSABSPQNNILMLGAVGAWDSGTGTVQKTPHGLIFSKQAFQIQLQDRNHSYLGYS 429
392 AAIIILNRVQSLVLCAPRYOHIGLVAMFRONTGWMESNANV-----KGTQICAYEGASL 445
430 VASISTGNSVHFVAGAPRANTYQGVLYSYN-----ENGNTVVIQSGDQIGSYFGLV 484
446 CSVDVDSNGSTDLVLGAPHYEOTR--GGQSVCPPLRGQARWQCDVAVLYGECQCPWG 503
485 CAVDVNKDTITDVLVAGPMYNDLAKKEGRVYLFITKG-ILNWH--QFLEPGNLENA 541
504 RFGAALTVLGDVNGDKLTDVAIGABGEDNRGAVYLFHGTSGSGISPSHSQRIAGS--XL 561
542 RFGSAIATLSDINMGDNFVIVGSPLENQNSGAVIYNGHEGM-IRLRYSKILGSDRAF 600
562 SPRLQYFGQSLGGQDLTMDGLVLTGAGCHVLLRLSOPVLAVKAIMFENPREVARNVF 621
601 SSHAQYFGRSLDYGDLNGDSITDVSAGFQVQVQLWSQSIADVSVDASTPKKI--TLL 658
622 ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQISVVTYDLALD-----SGRPHSRAVEN 677
659 NKAAEI-----KLKLCF-----SAKFRPTNQNNQVAIVYNIIDEDQFSRVSRLGFK 707
678 ETKNSTRQTOVLGLTQTCB--TLKQLPNCIEDPVPVILRLNFSL--VGTPLSAFGLN 733
708 ENNERCLQKTMVISAQQRCEYIIHQEPS---DIISPLNLCWNISLENPGT-----756
734 RPLVLAEDQRLFTALPFKEKNCNDNICQDLDITP-----SPMSLCLVWGGPRENVTV 789
757 NPALAEAYSETVKVFSIPFHKDCGDGVCISDLVNLVQQLPATQOQPPFIIVSNQNKRLTFSV 816
790 TVRNDGDSYRTOVTFPPFLDLSYRKVSTLQNRQSORSLRACESASST-EVSGALKSTS 848
817 OLKKNKESAYNTELVDFSENLF-----ASWMPVDGTEVTQCIASSOKSVT 864
849 CSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFOLELPVKY 908
865 CNVGPALKSKQQTFTINFDFNLQ-NLQNAQSIAPRALSESQENMADNSVNLKLSLLY 923
909 AVYMWVTSHGVSUKYLNFTASNTSRVMOHQYQVSNLQOR-----SLPISLVFLV 958
924 DAEIHIT-RSTNINFEVSLDGNVSVV-HSFE--DIGPKFIFSIVKTVTGSVPVMSA---976
959 PVLNQTVIWDROPVTFSEN---LSSSTCHTKE-----RLPSHSDFLAE-998
977 -----SVIITHIPOYTKDKPLMYLTGVHTDQAGDISCEAEINPLKIGQTSVSPKSEN 1030
999 LRKAPVNCISAVCQRIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEI-1057
1031 FRHIKELNCRATSCSNIMCWLRLDQVKGEYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
1058 LFNDSVFTL-----LPGQAFVRSQTEKVEPPE-VPNPLPLIVGSSVCGLLILLALITA 1110
1091 TYNPQIYVIBENTVITP-----LTIMKPEKVEVPTGVIVGSIAGIILLALVA 1140
1111 ALVKLGFFKRYQKDM 1125
1141 ILWKLGFFKRYKEXM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
P:169-344/Domain: von Willebrand factor type A repeat homology <WVA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;

Best Local Similarity 27.8%; Pred. No. 3.1e-65;

Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;

QY 1 FNLDTENAMTFQ-ENARGFGOSVVLQ---QGSRVVVGAPQEIIVANQSGSLYQC---DYST 54
DB 27 YNVGLPGAKIFSGPSSSEQGYVSQQLTNPQGNWLLVGSFPMGDMGVTKCPVDLPT 86
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPPPQLLACGPTVHOTCSSENTYVKGIC 107
DB 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCGLPLWAHQCGNQYATGIC 146
QY 108 FLFGNLRQOPQ---KPFELRGCPQEDSDIAFLIDGSGSIIPHFRMKKWWSTVMEOL 164
DB 147 ----SDVSPDFQFLTFSFPAVQACPSL-VDVVVVWCDENSIYP--WEAVKNFLVKFVTGL 199
QY 165 K--KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLG-RTHATATGLRKVVRE 221
DB 200 DIGPKTKQVALIOYANEPRIIIFNLNDFETKEDMVOATSETROHGGDLTNTFRAIEFADY 259
QY 222 LFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDREGVIRYVIGV-----GDAFR 276
DB 260 AYSQTSGGPAGATKVVVVYTDGESH-DGSKLTVIQCCNDDEILREGIAVLGNRLNLD 318
QY 277 SEKSRQELMTVASKPRDRHVFOINNFEAKTIONQLREKIFAIEGTQTGSSSSFEHMSQ 336
DB 319 TKNLIKEIKAIATASTPTERYFFNVVADEAALLEKAGTLGEQIFSGIEGTQVQ-GDNFQEMAQ 377
QY 337 EGESA--AITSNGPLLSTVGSVDWAGGVPLYTSKEKSTFINMT--RVSDMN-DAYLGYA 391
DB 378 VGFSADYAPQNDILMLGAVGAFDWSGLTQVETSHKPVIPFKAQFDQVLQDRNHSFLGYS 437
QY 392 AAIIILNRVQSLVLCAPRYOHIGLVAMFRONTGWMESNANV---KGTQICAYEGASLCS 447
DB 438 VAAISTEDGVHVFAGAPRANTYQGVLYSYN---QGNVTVIQSHRGDQIGSYFGLVCS 494
QY 448 VDVDSNGSTDLVLGAPHYEOTR--GGQSVCPPLRGQARWQCDVAVLYGECQCPWGFR 505
DB 495 VDVQDQITDVLVAGAPTMYNDLAKKEGRVYLFITKGLNQHQ---FLEPGEGTGNARF 551
QY 506 GAALTIVLGDVNGDKLTDVAIGABGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564
DB 552 GSAIAALSINMGDNFVIVGSPLENQNSGAVIYNGHEGT-IRTKYSQKILGSDNGAFRR 610
QY 565 -LQYFGQSLGGQDLTMDGLVLTGAGCHVLLRLSOPVLAVKAIMFENPREVARNVFEC 623
DB 611 HLOPFORSLDGYDGLNGDSITDVSIGALGOVQLWSQSIADVAIEALFTP-----660
QY 624 NDQVVKGEAGEVRVCLHVQKSTRDLREGQISVVTYDLALD-----SGRPHSRAVENET 679
DB 661 -DKITLLNKDAKITLCLCFRABFRPAGQNNQV--AILFNMTLDADGHSRVRTSRGFREN 717
QY 680 KNSTRQTOVLGLTQTCET--LKLQPLNCIEDPVPVILRLNFSLVTPLSAFGLNRPVL 737
DB 718 SERFLQNMVVADEVQKSEHHISIQKPS---DVNPLDLRLVDISLENPGTS-----PAL 768
QY 738 AEDAQRLLTALFPFEKNCNDNICQDLDLSI-----TFSPMSLCLVWGGPRENVNT 788
DB 769 EAYSETVKVFSIPFYKECGSDGICISDLIDVQQLPAIQTSF-----IVSNQNKRLTFS 823
QY 789 VTVRNDGDSYRTOVTFPPFLDLSYRKVSTLQNRQSORSLRACESASST-EVSGALKST 847
DB 824 VILKNRGESAYNTVLAEFSENLF-----ASFMPVDQTEVTCEVSGSQSV 871
QY 848 SCINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFOLELP 905
DB 872 TCDVGYPALKEQOQVTFITDFNLQ-NLQNAQAIINQAFSESQ--ETNKAADNSVSLTIP 928
QY 906 VKYAVTVMTVSHGVSTKYLNFTASNTSRVMOHQYQVSNLQOR-----SLPISLV 955

436	LGTYVNSATIPGDVLVYTAGQPRYNHTGQVVIYKMGEDGNINILQTLGGBEIQISYFGSVLTT	495
448	VDVDSNGSTDLVLVIGAPHY-----YEQTR-GQSVSVCPPLRQQRARWQCD	491
496	IDIKDYSYTDLLLVGAPMYMGTEKEBOQKVVIYAVNQTRFYEQMGLEPIRQTCSSLKON	555
492	AVLYGEOQPMWG-RFGAALTVLGDVNGDKLTDVAIGAPEEDNRGAVLYFHTGTSQGISP	550
556	SCTKENKNEPCGARFGTAIAVAKDLNVDFNDVVIAGAPLEDDHAGAVIYHG-SGKTIRE	614
551	SHSQRISAGSKLSPLROYGQSISGGDLTMQGLVDLTVGAQGHVLLLRSQPVLRVKAI	610
615	AYAQRIPSGGGCKTLKFFGQSIHGEMDLNGDGLTDVTI GGLGGAALFWARDVAVVKTWN	674
611	FNPREVARNVFECNDQVVYKBEAG--EVRVCLHVQ-KSTRDLRBEQIQSVVYTYDIALDS	667
675	PEPNKVIQKNCNCR---VEGKETVCINATMCFHVFLKSKEDSIYEADLQ----YRVTLDS	727

Db	728	LQI	S	R	S	R	F	S	G	T	Q	E	R	K	I	O	R	N	I	T	V	R	S	E	---	C	I	R	H	S	F	M	L	D	K	H	D	F	Q	770																		
Qy	713	P	I	V	L	R	N	P	S	L	V	G	T	P	L	S	A	F	G	N	I	R	P	V	A	E	D	A	Q	R	L	F	T	A	L	P	F	P	K	N	G	N	D	I	C	O	D	D	L	S	I	T	F	S	M	772		
Db	771	S	V	R	V	L	D	F	N	L	T	-	D	P	E	N	G	---	P	V	L	D	A	L	P	N	S	V	H	E	H	I	P	F	A	K	D	C	G	N	K	E	R	C	I	S	O	L	T	L	N	V	S	T	824			
Qy	773	S	L	D	C	L	V	G	G	P	R	-	F	N	V	T	V	R	N	D	G	S	V	R	T	O	T	F	F	F	P	L	D	L	S	Y	K	V	S	T	L	Q	N	R	S	Q	R	S	W	L	A	831						
Db	825	E	K	S	L	L	I	V	K	S	O	H	D	K	F	N	S	L	T	V	K	N	G	D	S	A	N	T	R	T	V	Q	H	S	P	N	L	I	F	S	G	E	E	I	O	K	D	---	S	876								
Qy	832	C	E	S	A	S	S	T	E	V	S	A	L	K	S	T	C	S	I	N	H	P	F	P	E	N	S	E	V	T	F	N	I	T	E	D	V	D	S	K	A	S	L	G	N	K	L	-	L	K	A	N	V	T	S	E	N	890
Db	877	C	E	S	N	---	---	---	Q	N	I	T	C	R	V	G	F	P	F	A	G	E	T	V	T	F	K	I	F	O	P	N	T	S	H	L	E	N	A	I	H	L	S	A	T	S	D	S	E	927								
Qy	891	N	M	P	R	T	K	T	E	F	O	L	E	P	V	K	Y	A	V	---	Y	M	V	T	S	H	G	V	S	T	---	---	K	Y	L	N	F	T	A	S	E	N	T	R	V	M	Q	H	Q	939								
Db	928	P	L	S	L	N	D	N	E	V	N	I	S	I	P	V	K	E	V	G	L	Q	F	Y	S	S	A	S	E	H	I	S	A	A	N	E	T	I	P	E	F	I	N	S	T	---	E	D	I	G	N	E	I	N	V	985		

[illegible]

RESULT 11

integrin alpha-2 chain precursor - human
A33998
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor
C:Species: Homo sapiens (man)
C>Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 21-Jul-2000
C:Accession: A33998; B56793; A53117
B:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha (2) -subunit
A:Reference number: A33998; PMID:89108879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAK>

A;Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catalmel, B.; Parmentier, S.; Leung, L.B.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GP1a, GP1c, GP1c*, GP1a and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:L28031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105, 112, 343, 432, 460, 475, 699, 1057, 1074, 1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.7%; Pred. No. 5.1e-65;
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDTENAMTQ-ENARGFGOSVQL---QGSRVVVGAPQEIIVAAQNGSLYQC--DYST 54
Db 30 YNVLGPEAKIFSGSSSEQFGVAVQFIPNKGWLLVSGSPGFENRGDYKCPDLST 89
Qy 55 GSCEPIRLQ-----VPEAVNMSLGLSLAATSPQALLACGPTVHTQTSNTYVKGLC 107
Db 90 ATCEKLNQTSIPNVTETKNTNMSLGLILTRNMGTFGLTCGLPWAQCCQNGYTTGVC 149
Qy 108 FLFGSNLRQOQKPEALRGCPQSDSDIAFLIDSGSIIIPDFRMEKWTSTWEOQLK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPL-IDVVVVCDESNSIYPWD--AVKNFLEKFFVOGLDIG 205
Qy 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPIITQLLG-RTHATGLRKVVRELFN 224
Db 206 PTKTQVGLIQANNPRVFNLTNYKTEEMIVATQTSQYGGDLTNTFGAIQYARKAYS 265
Qy 225 ITNGARKNAKILFLITDGEKFGDPLGYEDVIPELDREGVIRYVIGV-----GDAPRSEK 279
Db 266 AASGRRSATKVMVVVTDGESH-DGSMKAVIDQCNHNDILRFGIAGVLGNRNLALDTKN 324
Qy 280 SRQELNTVASPPRDHVPQINNFEALQIKIQAIGTQTGSSSSSEHEMSQEGF 339
Db 325 LIKEIKATASIPTEFYFNVDSEAALEKAGTLGEQIFSIETGQV-GDNFQMSQVGF 383
Qy 340 SAATISNGP--LLSTVGSYDAGGVFLVTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394
Db 384 SADYSSQNDILMLGAVGAFSGTIVQKTSHGHLIFPKQAFDQILQDRNHSYLYGVAA 443
Qy 395 ILNRNVQSLVGLGAPRYQHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFVACPRANTVQGVLYSVN-----ENGNITVIAHRGDDQIGSYFGSVLCSV 498
Qy 449 DVDSNGSTDLVLIGAPHYEQTR--GGQVSVCLPRGQARWQCDVLYGQGGQWPRFG 506
Db 499 DVDKDTITDVLVAGPMVMSDLKBEGRVYLTIKKILGQHQ---FLEGPEGIENTRFG 555
Qy 507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVVFLPHTGSGSGISPSHSQRIAGS--KLSPR 564

Db 556 SAIAALSNDIMDFNDVIVGSPLENQNSGAVIYNHQQT-IRTKYSKILGSDGAPRSH 614
Qy 565 LOYFGQSLGGODLTMDGLVDLTVCAGQGVLLLRQPVLRVKAIMEENPREVARNVFECN 624
Db 615 LOYFGSLDGYGDLNGDSITDVISGAGGVQVQLMSQSTADVAIEASFTEPKI--TLVNKN 672
Qy 625 DQVWKGKEAGEVRVCLHVQKSTRDLRREGIQSVVYTYDLALD-----SGRPHSRVAFNETK 680
Db 673 AQII-----LKLCF---SAKPRTKQNNQVAIYNNITLDADGFSRVRTSRGLFKENN 721
Qy 681 NSTRQQTQVLGUTQTC--ETLKLQIPNCIEDPVSFIVLRNLSLVTGTPLSAAGNLRPVLA 738
Db 722 ERLCKQNVVNAQSCPEHIIYIQBPS---DVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFPFEKNCNDNICQDLSITF---SFMSLDCLVVGGPREFNVTVTRND 794
Db 773 AYSETAKVSIPIPHKDCGEGDGLICISDLVDVQIIPAAQEQPIFVSNQKRLFSVTLKNK 832
Qy 795 GEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLPF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
Qy 854 PIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEFQLELPVKYAVMV 913
Db 881 PALKRQQVTFITFDNLQ-NLQNASLSFOALSESQEEKADNLVNLKIPULLYDAEI- 938
Qy 914 VTSHGVS TKYLNFTASENTSRVMQHOYQVSNLQOR-----SLPISLVFLV---- 958
Db 939 ---HLTRSTNINPFISSDGNVPSIVHSFEDVGPKEFISLKVTTGSGVPVSMATVIIHQ 995
Qy 959 -----PVRINQTVIWDPRQVTF-SENLSSTCHTKER 988
Db 996 YTKENKMLVLTGVTQDKAGDISCNADINPLKIGQT---SSSVSPKSENFR---HTKE- 1047
Qy 989 LPSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIOEENATLKNLSFDWYIKTSHNH 1048
Db 1048 -----LNCRTASCNVTCWLKDVHMKGBYFVNVVTVTRINWGTFFASSTFQT 1091
Qy 1049 LLIVSTAEL-LFNDVSTELLPGQAFVRSQETKVEPFEPNP-----LP--LIVGSSVG 1100
Db 1092 VQUTAAEINTYNEIYVI-----EDNTVTIPLIMKPDKEAEVPTGVIGSIIA 1141
Qy 1101 GLLLALITALYKLGFFKRYQKDM 1125
Db 1142 GILLALLVALKLGFFKRYQKDM 1166

RESULT 12

A41131

Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N;Alternate names: integrin alpha-4

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999

C;Accession: A41131; S16742

J;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzman, B.; Weissman, I.L.

J. Cell Biol. 115, 1149-1158, 1991

A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte

A;Reference number: A41131; MUID:92064645; PMID:1840602

A;Accession: A41131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1039 <NEU>

A;Cross-references: EMBL:X51176; NID:g51484; PIDN:CAA37316.1; PID:g51485

C;Superfamily: integrin alpha-4 chain

C;Keywords: cytoskeleton; transmembrane protein

Query Match

Best Local Similarity 11.3%; Score 663; DB 2; Length 1039;

Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

Qy 1 FNLDTENAMTQ-ENARGFGOSV-VLOGSR--VVGAPQEIIVAN-----QRGLSYQCDY 52

Db 41 YNLDPENALLYQGPSGLTFGYSVVLHSHGSKRWLIIVGAPTASWLSNVAIVRCGI 100
QY 53 STG---SCPIRLQVP-----VEAVNMSLGLSLAATTSP-PQLLACG---PTVHQ 95
Db 101 RKNPNQTCQLQSGSGSGPCGKTCLEERDNQWLVTLRQPGENGSIYTCGRWKNVIFY 160
QY 96 TCSNTYVYKGLCLFGLFNSLRQQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKE 155
Db 161 MKSDNKLPTGICYVMPSDLRTLSK-----185
QY 156 WSTVMEQLKSKTFLSLMQYSEPRHFHTFKFQNNPNRSLIKPITQLLGRTHATGL 215
Db 186 -----RMAPCYKDYT-----195
QY 216 RKVVRLEFNITNGARKNAFKILLTDGEKFGDPLGVEDVIPELDREGVIRYVIGVDAP 275
Db 196 -----RKFGZ-----200
QY 276 RSEKSRQELNTVASKPPRHVFOINNFEALKTIQNLREKIFAIEGTQTGSSSFBHEMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGFSAATNSGPLLSTVSGSYDNAGGVFLY--TSKEKSTFINMTRVDSMDNAYLGYA-- 391
Db 206 QAGISSFYTDLIWVGAPSSYWTGTVFVYNTTQYKAFVD--RQNVKFGSYLGSYVG 263
QY 392 AAILRNRYQSLVGLGAPRYGHICGLVAMFRQNTQMESNANV-----KGTOLGAVFGSLCS 447
Db 264 AGHFRSPHTTEVVGAPQHQQIKAYIF-----SIDENELNIVEMKGLGSLYFGASVCA 319
QY 448 VDVSNGSTDLVLIGAPHYYEQRGGQVSQVCLPRGQRA--RMQCDVAVLXGEOQPMGRFG 506
Db 320 VDLNAGFSDL-LVGAPMQSTIREEGRVFY--INSGMGAVMVEMERVLVGSDKYA-ARFG 376
QY 507 AALTVLGVNKGDKLTVAIGAPDEDRNARGVYLPHGTSGSGISPSHSQRIAGSKLSPRIQ 566
Db 377 ESTANLGIDINDGFDIAGQEDDLRGAVIYNGRV-DGISSTYSQRIEQGQISKSLR 435
QY 567 YFGQSLSGGODLTMGDLVLTGCA--QGHVLLRSOPVLVRVKAIMFENPREVARNVFEGN 624
Db 436 MFGQSLSGQIDANNGVQDVAVAFQSDSAVLLTRPVVIVEASLS-HPESVNRTRKFDCT 494
QY 625 QVVVKGKEAGEVRVCLHVQKSTRDLREGIQISVWVYDLALDSGR---PHSRAVF--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGVPGYIVLVFNVSLLVHRKAEPSRPFYFNGT 547
QY 680 KNSRTRQTVLGTQTCETLKLQLPNCIEDVPSPVILRLNFSL-----VOTPLSAFNGLR 734
Db 548 SDVITGSIRVSSSGEKRCRTHQAFMRKQVDRDILTPHVEATYHLGHHVITKRNTTEPPPLQ 607
QY 735 PVLAEDAQR-LFTALPFEKNCNGDNICODDLSITFS-----FMSLDCLVVGCPREFNV 787
Db 608 PILQOQKEDVIRKMINARFCAYEN-CSADLQVSAKVGLPKPYENKTYLVAGSMKTIWL 666
QY 788 TVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACASASTEVSGALKST 847
Db 667 NYSFLNAGDDAYETLLNVQLPTGLYFIKILDEEK-----QINCE---VTSSGIVK-L 716
QY 848 SCISNHPPIFPENSEVTNITFDVDSKASLGNKLLKANVTSEN--NMPRTNKTETFOLELP 905
Db 717 ACSGLGIYVDRLSRIDISFLDDVSSLSRAHEDLSISVHASCENEGELDQVRDRNRVLTIP 776
QY 906 VKYAVVMVTSHV--STKYNLTASNTSRVMQHQ-----YOVSNLQORSLP-ISLVEL 957
Db 777 LRYEV--MLTVHGLVNPPTGFVYGSSENEPETCMAEKLNLTHVINTGISMAFNVSVKIM 834
QY 958 VPVRLNOTVWDRP--QVTFSENLSTCTHKE-----RLPSSHDDFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLDOVTTGQCHFKYHGRECTPAQKGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVAVNCSTIA--VCQRIQCDIPFFGQIEEFNATLKGNLSPDWYIKTSHNHLITVSTAEIL 1058
Db 892 KR-LLYCMKADQHCFLDCLN---FGWMESG-----KEASVHIQLEGPSIL 933

QY 1059 FNDVSFTLLPGQGAFAVRSQTETKVFPEFVNP-----1090
Db 934 ENDESSL-----KFEIKATAFPEPHKVIELNKDENVAHVFLGLELHHQPKRHF 983
QY 1091 -LPLIVGVSUGQLLLALITAAALYKLGFFKQYKMMSE 1128
Db 984 TIIITITISLLGLVILLISCVMMKAGFFKQYKSILOE 1022

RESULT 13

S06046

N.Integrin alpha-4 chain precursor - human

C.Species: Homo sapiens (man)

C.Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999

C.Accession: S06046; A39355; D28018

R.Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.

EMBO J. 8, 1361-1368, 1989

A.Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other int

A.Reference number: S06046; MUID:89356603; PMID:2788572

A.Accession: S06046

A.Molecule type: mRNA

A.Residues: 1-1038 <TA>

A.Cross-references: GB:X16983; EMBL:X15356; NID:G33945; PIDN:CAA34852.1; PID:G33946

R.Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.

Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991

A.Title: Characterization of the alpha-4 integrin gene promoter.

A.Reference number: A39355; MUID:91239513; PMID:2034655

A.Accession: A39355

A.Molecule type: DNA

A.Residues: 1-93 <ROS>

R.Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A.Title: The very late antigen family of heterodimers is part of a superfamily of mole

A.Reference number: A94151; MUID:87204112; PMID:3033641

A.Accession: D28018

A.Molecule type: protein

A.Residues: 40-50, E' 52-53 <TA2>

C.Genetics:

A.Gene: GDB:ITGA4; CD49D

A.Cross-references: GDB:128032; OMIM:192975

A.Map position: 2q31-2q32

C.Superfamily: integrin alpha-4 chain

C.Keywords: Cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans

F;1-39/Domain: signal sequence #status predicted <SIG>

F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.9%; Score 642; DB 2; Length 1038;

Best Local Similarity 23.8%; Pred No. 2,8e-16;

Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;

QY 1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVVGAPEIIVAAAN---QRGSLYQCDY 52

Db 40 YNVDTESALLYQGPHTLFGYSVVLHSHGCAWLLVGAPTANWLANASVINPGAIYRCRI 99

QY 53 STG---SCPIRLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHTQTCSENYVKGCLFL 109

Db 100 GKNPGQTCQLQGLSP-----NCEP---CG---KTCLEERDNQWL---133

QY 110 FGSNLRQOPKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKT 169

Db 134 -GTVLSRQGE-----NGSIVTCGHR-----W---KN 156

QY 170 LFSLMQYSEPRHFHTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGA 229

Db 157 IFVI-----160

QY 230 RKNAPKILFLTDGSKFGDPLGVEDVIPELDREGVIRYVIGVDAPFRSEKSRQELNTVAS 289

Db 161 -KNENK---LPTGCGYGP-----PDL-----RTELSKRIA 187


```
Qy 290 KPPRDHVFQI--NNFEALKTIONLREKIFAETGTGTGSSSSFEHMSQEGFSAAITSNP 348
Db 188 PCYQDYVKKGFENAS-----CQAGISSFTYKDLI 217
Qy 349 LLSTVGSVDWAGVFLY--TSKSKSTFINMTRVDSMDNDAYLGYA--AAIILNRVQSL 403
Db 218 VMGAPGSSYMTGSLFVNTITNKYKAFDKQNKQVF--GSYLGSYGAGHFRSQHTTEV 274
Qy 404 VLGAAPRYOHLGLVAMFRQNTGMWESNANV---KGTGIGAVFGASLCSVDVDSNGSDLV 459
Db 275 VGGAPQHEQIGKAVIF---SIDEKELNILEMKGKGLGSYFGASVCAVDLNAUGFSDL- 329
Qy 460 LIGAPHYYEOTRGQGVSCPLPRGORARWQC-DAVLYGEOQGPGRFGAALTVLGDVNGD 518
Db 330 LVGAPMQSTIREGRVFY--INSGSAVMNAMETNLVGSOKYA-ARFGESIVNLGDDND 387
Qy 519 KLTVAIGAPEEDNRGAIVLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDL 578
Db 388 GFEDVAIGAPQEDDLOGAIIYINGRA-DGISSTFSQRIEGLQISKLSMFQSGISQIDA 446
Qy 579 TMDGLVDLTGCA--QGHVLLLRSPVLVRVKAIMEFNPREVARNVFECNDQVVGKEAGEV 636
Db 447 DNGYVDVAVAGFASDGAVALTRPVIVDASLS-HPESVNRTRKFDG-----VENGWP 498
Qy 637 RVCLHVQKSTRDLREGIQSVVTVYDLALDSGR-----PHSRVAFNETKSTRRTQTVLG 691
Db 499 SVCIDLTLTCSYKKEVPGYIVLFYNMSLDVNRKAESPFRPYFSSNGTSDVITSGIQVSS 558
Qy 692 LTQTCETLKLQLPNCIEDPVSPIVLRNFSLVGTPLS-----AFGNLRPVLAEDAQR-LF 745
Db 559 REANCRTHQAFMRKDVRLDITPQIETAAHYLGHPIVSKIRSTEETPPLOPILQKKEKDIM 618
Qy 746 TALPFEKNCNDNICQDDLSIT--FSPM-----SLDCLVCGPREFNVTVVRNDGDSY 799
Db 619 KKTINFARFAHEN-CSADLOVSAKIGFLKPHENKTYLAVGSMKTLMLNLSLFWAGDDAY 677
Qy 800 RTQVTFPPDLRYKRVSTLQNRQSRWSRLACESASTEVSGALKSTCSINHPIPPEN 859
Db 678 ETTLHVKLVPGLYFIKILEBEK-----QINCE---VTDNSGVVQ-LDCSIGIYVDHL 727
Qy 860 SEVTNFTTFDVSQKASLGKLLKANVTSEN--NMPRTNKTFOLELPVKYAVVMVTSH 917
Db 728 SRIDISFLDVSLSRAEDLSITVHATCENEEMDNLKHGRVTVVAIPLKYEVKLTVRHF 787
Qy 918 GVSTKYNLFTASENTRV----MOHOYQVNLGORSIP-LSLVFLPVRLNQTVINDRP 971
Db 788 VNPTSFVGSNDENEPETCMVEKMNLTTPHVINTGNSMAPNVSVIEMVNSFS-----P 840
Qy 972 QVTFSENL-----SSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIPFFGIQ 1025
Db 841 QTDKLFNILDVQTITGCHFENYQ-----RVCALEQ-----Q 872
Qy 1026 EEFNATLKGNSLDFWYIKTSHNHLIVSTAE-----ILFN-----DSVFTLLPQGG 1071
Db 873 KSAMQTLKGIYVRF---LSKTDKRLLYCIKADPHCLNFCNFGKMSGKEASVHIQLEGRP 929
Qy 1072 AFVRSQRTKVEPEEV-----PNP-----LPLIV 1095
Db 930 S-ILEMDETALKFEIRATGPEPNRVIELNKDENVAHVLLGLHQRPKRYFTTIVIS 988
Qy 1096 GSSVGGLLLLALITAAALYKGLFFKRYQKDMSE 1128
Db 989 SLLGLLGLIVLLLSIYVMWKAGFFKRYKSILOE 1021
```

RESULT 14
158409
Integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994

A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in em.
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
A:Cross-references: GB:D25303; NID:G464180; PIDN:BAA04984.1; PID:G533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel par
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:L24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 1.9e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

```
Qy 331 EHEMSQEGFSAAITSNGLPLSTVGSYDWAGGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389
Db 190 EHSGCQAGIAGFFTEELVVMGAPGSGFYWAGTIKVLNLT-D-NLYLKLNDVIMNRRYTYLG 248
Qy 390 YA--AAIILNRVQSLVGLGAPRYQHIGLVAMFR--QNTGMWESNANVKTQIGAYFASL 445
Db 249 YAVTAGHFSPSTIDVVGGAPODKGIGKVIYFRADRRSGTLIKIFQASGKKMGSYFGSSL 308
Qy 446 CSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPLPRGORARWQCDAVLYGEOQGPWGRF 505
Db 309 CAVDLNGDGLSDL-LVGAPMFSIRDEGQTVY-INRGNGALEE-QLALTGD-CAYNAHF 364
Qy 506 GAALTVLGDVNGDKLTDVAIGAPEEDNRGAIVLFHGTSGSGISPSHSQRIAGSKLSPRL 565
Db 365 GESIASLDDLNDGDFDVAIGAPKEDDFAGAVYIYHGDAG-GIVPQYSMKLSGQKINPVL 423
Qy 566 QYFGQSLGSGQDLTMDGLVDLTGCA--QGHVLLLRSPVLVRVKAIMEFNPREVARNVFEC 623
Db 424 RMFGQSSIGGIDMGNGYPODVTVGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
Qy 624 NDQVVGKEAGEVRV-CLHYQ-----KSTRDLREGQIQSVVTVYDLALDSGRPHSRV 675
Db 483 HD-----GQPVNCLNVTTCFSPHGKHVPEI---GLNYVLMAVVAKEKQMPRVY 531
Qy 676 F---NETKNSTRRTQTVLGTQTCETLKLQLPNCIEDPVSPIVLRNLSL-----VGTPLS 728
Db 532 FVLLGETMGQVTEKLQLTVMETCRHYVAHVKKRVQDVISPIVPEAAYSLSHVTVGBER 591
Qy 729 AFGNLRPVL-----AEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFM--SLD-----CL 777
Db 592 ELPLLTPLVRWKKGQKIAQKNQTV---FERNCRSED-CAADLQLOGKLLSSMDKLTLYL 647
Qy 778 VVGSPREFNVTVVRNDGEDSYRTQVTFPPDLDSLYKRVSTLQNRQSRWSRLACESASS 837
Db 648 ALGAVKNIISNISINLGGDAYDANVSFNVSRELPF-----INWQKEEMGISCELLES 701
Qy 838 TEVSGALKSTCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNK 897
Db 702 DFL-----KCSVGFPFMRKSKYEFVIFDTSHLS--GEEVLSPFVTTAQSG--NTER 750
Qy 898 TE-----FOLELPVKYAVVMVTIS-----HGVSTKYNLFTASENLS-----RVMOHQY 940
Db 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFFVYGESVDAANFIQLDLECHFPQINITL 810
Qy 941 QVSNLGRSIPISLVFL-VPVRLN-----QTVMWRPQVTVSENLSSTCHTKER 988
Db 811 QVNTGTSTLPGSSVSISFPNRLSSGGAEMFHVQMVVGQSKGNCFSQKNPTPCLIIQEQ 870
Qy 989 LPSHSDFLAELURKA-----PVVNCISIAVCORIQCIPFFGIQEBFNATLKGNSLSD 1039
```


Db .371 ENIFHTIFAFPTKSGRKVLDCBKGISCLTAHCN-----PSALAKESRTI-----D 917
Qy 1040 WYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTETKEP----- 1084
Db 918 IY-----MLNT-EILKDDSSVIO-----FMSRAKVVDPALRVEIAHGNPEEV 962
Qy 1085 ----FEVFN---PLPLIVG-----SSVGGLLLLALTAALYKLGFPKQYKMM 1126
Db 963 TVVFEALHNLPRGVYVGVGIIAISLLVGLILPLLALLVLMKMGFFRRYKEII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.4%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 2,3e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 336 QEGFSAATSNCP--LLSTVGSYDAGGVFLYTSKEKSTFINMTRVDS-----DMNDAYLG 389
Db 181 QAGFSGIIFSDNSALVMGAPGSYYLQGGYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
Qy 390 YAAAI--ILNRVQSLVLCAPRYQHI-GLVAMFRONTGMWESNANVKGTGICAYEGASLC 446
Db 240 YSLALGDFNGDGVQYVVGTPRAESLMGLVAIFDNLNQFN---QVMGTQIVAYFGYSVT 296
Qy 447 SYVDVDSNGSTDILVLCAPHYEQTRGGQVSVCPLRGQARMCQDAVLYGEO----- 498
Db 297 VVDI-NNDTYDILLVGAPEMDGPAIQ-----RWEAGAVVYLLQNPDPVGPQA 343
Qy 499 -----GQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGS 546
Db 344 SNRLSLSSTLIGGQIRSRPGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIHG-SAN 402
Qy 547 GISPSHSQRIAGSKLS-PRLOYFGOSLGGQDLTWDGLVDLTVGAQ--CHVLLLRSPVL 603
Db 403 GLKSTPAQVLTSTLGHSGITTFGFSLOGQDMKKNYPDLLVGAESANTAVLIIRPVV 462
Qy 604 RVKALMEFNPREVARNFECDNVQVKGEAGEVRVCLHVQKSTRDRLEGOIQSV----- 658
Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTWVTSFIAMT 497
Qy 659 -----VYDIALDSG-RPHSRVFNENKSTRQTVLGL-TQTCETLKLQ 702
Db 498 CFTYTGNYLPDHDIDISYTVVDSGIIANRRAMFVNDMNSIETKTRRLAVSTQFCPLRAY 557
Qy 703 LPNCIEDPVPVILNLSLVGTPLSAFN-----LRPLAEDAQRLEFALFPFEK 753
Db 558 VGNSEDKLTPIKVTLOYDL-----NNDESRLQPHIILPDIIDMATMSTQTKQVSIQN 609
Qy 754 NCGNDNICODDLSITFSFMSLDCLVVGPREFNVTVVRNDCGEDSYRTQVTFPFFPLDLSY 813
Db 610 NCVN-NICLPDLDVTV-PNLNIVIGTQELTLDVSNRRGEDAFQSLSVYPLGLQF 667
Qy 814 RKVSTLQNSQRSRLACESASSTEVSGALKSTSCSINHPIFPEN-----SEVTFNIT 867

Db 668 VRL-----ERKANMDFSVTCSSESD-----LRITCDTGNPMVGKNILFGLTJLSTFQVS 717
Qy 868 FDVDS-----KASLGNKLLKLLKANVTSENNMPRTNKTEFQLELPVKYAVYVMTSHGVSTK 922
Db 718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCCTLKLLSASYPEIVMYSTQED 775
Qy 923 YL-----NPTASENTRVMQHOYQV-----SNLQORSLSPIIS-----LVFLVPV 960
Db 776 YVVPFPFPAKNASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834
Qy 961 RLNQTVIWDPRQ-----VTFSENLSSSTCHTKERLPSHSDFLAELRKAPVV 1005
Db 835 MTEBGTVCQLTQGANPEGVKLEPSTRAKUSNSTTQVSGRRKRREPEVAEALQTDN--VI 892
Qy 1006 NCSIAVCQRIQCDIPFFGIQEEFNAT-----LKGNIISFDWYIKTSHNHLIVSTAEILF 1059
Db 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
Qy 1060 NDSVFTLLPGQAFVRSQTETKVE--PPEVNP-----LP----- 1092
Db 929 QKAVSELTPVVQATIASASAANKTIPYNIPLPRDFSDSTKASTLVTTTELVPVPTPIAW 988
Qy 1093 -LIVGSSVGGLLLLALITAAALYKLGFPKR-----QYKDMMS--EGGPP 1132
Db 989 WIIIVSVLGGIILLIILGLWKCGFFERKKPGEKEYAPVASADKOGPP 1038

Search completed: November 25, 2003, 14:22:00
Job time : 19.9579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5836.5	99.2	1152	1	ITAM_HUMAN
2	4460	75.8	1153	1	ITAM_MOUSE
3	3459	58.8	1163	1	ITAX_MOUSE
4	3401	57.8	1162	1	ITAD_HUMAN
5	1547.5	26.3	1170	1	ITAL_HUMAN
6	1527.5	26.0	1163	1	ITAL_MOUSE
7	1148.5	19.5	1167	1	ITAE_MOUSE
8	1140	19.4	1179	1	ITAE_HUMAN
9	1093.5	18.6	1151	1	ITAI_HUMAN
10	1084.5	18.4	1189	1	ITAH_HUMAN
11	1071	18.2	1170	1	ITAZ_BOVIN
12	1057	18.0	1178	1	ITAZ_MOUSE
13	1055	17.9	1180	1	ITAI_RAT
14	1054	17.9	1181	1	ITAG_HUMAN
15	1051.5	17.9	1167	1	ITAG_HUMAN
16	663	11.3	1039	1	ITAA_MOUSE
17	642	10.9	1038	1	ITAA_HUMAN
18	630	10.7	1035	1	ITAG_HUMAN
19	593.5	10.1	1032	1	ITAG_XENLA
20	571.5	9.7	1066	1	ITAC_CRISP
21	567.5	9.6	1053	1	ITAC_MOUSE
22	555.5	9.4	1053	1	ITAS_MOUSE
23	546.5	9.3	1034	1	ITAV_CHICK
24	545.5	9.3	1050	1	ITAS_XENLA
25	537.5	9.1	1130	1	ITAG_HUMAN
26	535	9.1	1044	1	ITAG_MOUSE
27	534.5	9.1	1072	1	ITAG_CHICK
28	532.5	9.0	1066	1	ITAS_HUMAN
29	532	9.0	1049	1	ITAS_HUMAN
30	526	8.9	1048	1	ITAV_HUMAN
31	517	8.8	1044	1	ITAB_CHICK
32	512.5	8.7	1091	1	ITAB_MOUSE
33	498	8.5	1179	1	ITAT_MOUSE

RESULT 1

ITAM_HUMAN	ID	ITAM_HUMAN	STANDARD	PRT	1152 AA.
34	494	8.4	1396	1	ITA2_DROME
35	491.5	8.4	1146	1	ITA1_DROME
36	490	8.3	1033	1	ITAB_MOUSE
37	489.5	8.3	1039	1	ITAB_HUMAN
38	489	8.3	1025	1	ITAB_HUMAN
39	486	8.3	126	1	ITAM_CAVPO
40	471	8.0	1181	1	ITA7_HUMAN
41	470	8.0	1106	1	ITA7_RAT
42	462	7.9	1226	1	PAT2_CAEEL
43	444.5	7.6	1139	1	INAI_CAEEL
44	427	7.3	1115	1	ITA3_DROME
45	385	6.5	1000	1	ITA5_DROME

ALIGNMENTS

AC	P11215;				
DT	01-JUL-1989	(Rel. 11, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)				
DE	(Neutrophil adherence receptor).				
GN	ITGAM OR CR3A OR CD11B.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88315033; PubMed=2457584;				
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;				
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";				
RL	J. Biol. Chem. 263:12403-12411(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88190151; PubMed=2833753;				
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;				
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88257215; PubMed=2454931;				
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;				
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor MO1 (complement receptor type 3).";				
RL	J. Cell Biol. 106:2153-2158(1988).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93123748; PubMed=8419480;				
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;				
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";				
RL	J. Immunol. 150:480-490(1993).				
RN	[5]				
RP	SEQUENCE OF 9-1153 FROM N.A.				
RX	MEDLINE=89098993; PubMed=2563162;				
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;				
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).				
RN	[6]				

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
RT conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation";
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tonich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment";
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03925; AAA59544.1; -
CC EMBL; M18044; AAA59491.1; -
CC EMBL; J04145; AAA59903.1; -
CC EMBL; S5227; AAB24821.1; -
CC EMBL; S52152; AAB24821.1; JOINED.
CC EMBL; S52153; AAB24821.1; JOINED.
CC EMBL; S52154; AAB24821.1; JOINED.
CC EMBL; S52155; AAB24821.1; JOINED.
CC EMBL; S52157; AAB24821.1; JOINED.
CC EMBL; S52159; AAB24821.1; JOINED.
CC EMBL; S52161; AAB24821.1; JOINED.
CC EMBL; S52164; AAB24821.1; JOINED.
CC EMBL; S52165; AAB24821.1; JOINED.
CC EMBL; S52167; AAB24821.1; JOINED.
CC EMBL; S52169; AAB24821.1; JOINED.
CC EMBL; S52170; AAB24821.1; JOINED.
CC EMBL; S52173; AAB24821.1; JOINED.
CC EMBL; S52174; AAB24821.1; JOINED.
CC EMBL; S52180; AAB24821.1; JOINED.
CC EMBL; S52181; AAB24821.1; JOINED.
CC EMBL; S52184; AAB24821.1; JOINED.
CC EMBL; S52189; AAB24821.1; JOINED.
CC EMBL; S52191; AAB24821.1; JOINED.
CC EMBL; S52192; AAB24821.1; JOINED.
CC EMBL; S52203; AAB24821.1; JOINED.
CC EMBL; S52212; AAB24821.1; JOINED.
CC EMBL; S52213; AAB24821.1; JOINED.
CC EMBL; S52216; AAB24821.1; JOINED.
CC EMBL; S52219; AAB24821.1; JOINED.
CC EMBL; S52220; AAB24821.1; JOINED.
CC EMBL; S52221; AAB24821.1; JOINED.
CC EMBL; S52222; AAB24821.1; JOINED.
CC EMBL; S52226; AAB24821.1; JOINED.
CC EMBL; M76724; AAA58410.1; -
CC EMBL; M84477; AAA51960.1; -
CC PIR; A31108; RWU1B.
CC PDB; 1A8X; 17-JUN-98.
CC PDB; 1BHQ; 18-NOV-98.
CC PDB; 1BHQ; 18-NOV-98.
CC PDB; 1IDN; 25-NOV-98.
CC PDB; 1IDO; 01-AUG-96.
CC PDB; 1JLM; 11-JAN-97.
CC PDB; 1MLU; 07-AUG-02.
CC Genew; HGNC:6149; ITGAM.
CC MIM; 120980; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.
CC SIGNAL 1 16 INTEGRIN ALPHA-M.
CC CHAIN 17 1152 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 17 1104 POTENTIAL.
CC TRANSMEM 1105 1128 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 1129 1152
CC REPEAT 31 84 FG-GAP 1.

Query Match		99.2%	Score 5836.5;	DB 1;	Length 1152;		
Best Local Similarity		99.1%	Pred. No. 0;				
Matches 1127;		Conservative	7;	Mismatches	2;	Indels	1;
						Gaps	1;
QY	1	FNLDTENAMTFQENARFGQSVVQLGSRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI	60				
DB	17	FNLDTENAMTFQENARFGQSVVQLGSRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCBPI	76				
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOPK	120				
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOPK	136				
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEQLKSKTLFSLMQSYSEEF	180				
DB	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEQLKSKTLFSLMQSYSEEF	196				
QY	181	RIHFTPEFQNNPNRSLKIPITQLLGRTHRTATGLRKVRVRELFNITNGARKNAFKILFL	240				
DB	197	RIHFTPEFQNNPNRSLKIPITQLLGRTHRTATGLRKVRVRELFNITNGARKNAFKILVVI	256				
QY	241	TDGEKFGDPLGYEDVPELDREGVIRVIVGVGDPAFSEKSRQELNTVASKPRDHDVFOIN	300				
DB	257	TDGEKFGDPLGYEDVPELDREGVIRVIVGVGDPAFSEKSRQELNTVASKPRDHDVFOIN	316				
QY	301	NFEALXTIQNLREKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDMAG	360				
DB	317	NFEALXTIQNLREKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDMAG	376				
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLVYAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR	420				
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLVYAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR	436				
QY	421	QNTGMWESNANVKGTQIGYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP	480				
DB	437	QNTGMWESNANVKGTQIGYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP	496				
QY	481	PRQARWQCDVLYGEQGPWGRFGAALTDLGVDVNGDKLTDVAIGAPGEEDNRGAVYLP	540				
DB	497	PRG-RARWQCDVLYGEQGPWGRFGAALTDLGVDVNGDKLTDVAIGAPGEEDNRGAVYLP	555				
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ	600				
DB	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ	615				
QY	601	PVLVRKAIIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLRREGIQSVVT	660				
DB	616	PVLVRKAIIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLRREGIQSVVT	675				
QY	661	YDLALDSGRPHSRAVNETKNSFRROTQVLGLTOTCETLKLQLPNCIEDPVSPVILRLNF	720				
DB	676	YDLALDSGRPHSRAVNETKNSFRROTQVLGLTOTCETLKLQLPNCIEDPVSPVILRLNF	735				
QY	721	SLVGTPLSAFGLNRPVLAEDAQLFTLFPFERNKCGNDNICDDLSITFSFMSLDCLVVG	780				
DB	736	SLVGTPLSAFGLNRPVLAEDAQLFTLFPFERNKCGNDNICDDLSITFSFMSLDCLVVG	795				
QY	781	GPREFNVTVRNDGSDSTRQTQVTFPFLDLSYRKVSTIQNRQSRQSWRLACESASSTEV	840				
DB	796	GPREFNVTVRNDGSDSTRQTQVTFPFLDLSYRKVSTIQNRQSRQSWRLACESASSTEV	855				
QY	841	SGALKSTSCSINHPIFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTFF	900				
DB	856	SGALKSTSCSINHPIFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTFF	915				
QY	901	QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVQHQVQVSNLQQRSLPISLVLFPV	960				
DB	916	QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVQHQVQVSNLQQRSLPISLVLFPV	975				
QY	961	RLNQTVIWRDPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP	1020				
DB	976	RLNQTVIWRDPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP	1035				

QY	1021	FFGIQEEFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDSVFTLLPQOGAFVRSQTET	1080
DB	1036	FFGIQEEFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDSVFTLLPQOGAFVRSQTET	1095
QY	1081	KVEPEFVNPLPLIIVGVSSVGGLLLLALITAAALYKLGFFKRQYKQKMMSEGGPGABPO	1137
DB	1096	KVEPEFVNPLPLIIVGVSSVGGLLLLALITAAALYKLGFFKRQYKQKMMSEGGPGABPO	1152
RESULT 2			
ITAM_MOUSE			
ID	ITAM_MOUSE	STANDARD;	PRT; 1153 AA.
AC	P05555;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).		
GN	ITGAM.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88312584; PubMed=3044779;		
RA	Pytela R.;		
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";		
RT			
RL	EMBO J. 7:1371-1378(1988).		
RN	[2]		
RP	SEQUENCE OF 11-45 FROM N.A.		
RC	STRAIN=BALB/C; TISSUE=Spleen;		
RX	MEDLINE=86287312; PubMed=2942940;		
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;		
RA	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";		
RT			
RT	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).		
RN	[3]		
RP	SEQUENCE OF 17-28.		
RX	MEDLINE=85188276; PubMed=3887182;		
RA	Springer T.A., Teplow D.B., Dreyer W.J.;		
RT	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";		
RT			
RL	Nature 314:540-542(1985).		
CC	-!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.		
CC	-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.		
CC	-!- SUBCELLULAR LOCATION: Type 1 membrane protein.		
CC	-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.		
CC	-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.		
CC	-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.		
CC	-!- SIMILARITY: Contains 1 VWFA domain.		
CC	-!- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EWBL; X07640; CAA30479.1; --
 EMBL; M14293; AAA39484.1; --
 PIR; S00551; S00551.
 HSSP; P11215; IABX.
 MGI; 96607; Itgcm.
 InterPro; IPR000413; Integrin_alpha.
 Pfam; PF01839; FG-GAP; 3.
 Pfam; PF00357; Integrin_A; 1.
 Pfam; PF00092; vwa; 1.
 PRINTS; PR01185; INTEGRA.
 PRINTS; PR00453; VWFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; VWA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Calcium; Repeat.
 SIGNAL 1 16
 CHAIN 17 1153
 DOMAIN 17 1105
 TRANSMEM 1106 1129
 DOMAIN 1130 1153
 REPEAT 31 84
 REPEAT 2 7
 DOMAIN 164 350
 REPEAT 337 400
 REPEAT 401 452
 REPEAT 454 515
 REPEAT 517 575
 REPEAT 580 632
 CA_BIND 465 473
 CA_BIND 529 537
 SITE 1132 1136
 DISULFID 66 73
 DISULFID 105 123
 DISULFID 654 711
 DISULFID 770 776
 DISULFID 999 1023
 DISULFID 1028 1033
 CARBOHYD 58 58
 CARBOHYD 86 86
 CARBOHYD 391 391
 CARBOHYD 696 696
 CARBOHYD 734 734
 CARBOHYD 772 772
 CARBOHYD 801 801
 CARBOHYD 881 881
 CARBOHYD 907 907
 CARBOHYD 941 941
 CARBOHYD 980 980
 CARBOHYD 994 994
 CARBOHYD 1022 1022
 CARBOHYD 1045 1045
 CARBOHYD 1051 1051
 CARBOHYD 1076 1076
 SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;

Query Match 75.8%; Score 4460; DB 1; Length 1153;
 Best Local Similarity 73.8%; Pred. No. 2.6e-297;
 Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

QY 1 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAQNRGLSYQCDYSTGSCPEI 60
 DB 17 FNLDTEHPTQENAKGFSQNVQLGGTSVVVAAPQEAQVNTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLGSLNRLOQPOK 120
 DB 77 PLOVPEAVNMSLGLSLAVSTVPQOLLACGPTVHOTCKNTYVNGLCYLFGLGSLNRLOQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEMVSTVMEQLKKSKTLFSLMQYSEBF 180
 DB 137 FPEALRGCPQEDSDIVFLIDGSGSINNIDFKMKPEFVSTVMEQFKSKTLFSLMQYSEBF 196
 QY 181 RIHFTKPEFQNNPNRSLIKPITOLLGRTHRTATGLRVKVVRELPNITNGARKNAFKTLFLL 240
 DB 197 RIHFTENDFKRNPSPRSHVSPIKQLNGRTKTASGIRKVVRELPHKNGARENAAKTLVVI 256
 QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDAPFRSEKSRSLNTVASKPRPDHVFQIN 300
 DB 257 TDGEKFGDPLGYEDVPEADRGVIRYVIGVGNFKNPQSRRELDITASKPAGEHVFQVD 316
 QY 301 NPEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLSTVSGVSDWAG 360
 DB 317 NPEALNTIQNLQEKIFAIEGTQTGSSSFEHEMSQEGFSAITSNGLPGSVGSDWAG 376
 QY 361 GUPLYTSKESKTFINMTRVDSMDNDAYLGVAAILNRVQSLVLCAPRYQHIHGLVAMER 420
 DB 377 GAFLYTSKOKVTFINTTRVDSMDNDAYLGVASAVILNRVQSLVLCAPRYQHIHGLVAMER 436
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 480
 DB 437 ENFGTWEPTSIKSGIISYFGASLCSVDMDADGNTNLILIGAPHYYEKTRGGQSVSCPL 496
 QY 481 PRQARWQCDVLYGEGQGPWGRFGAALTGLGVNGDKLTDAIAGPEEDNRGAYVLF 540
 DB 497 PRG-RARWQCEALLHGDQGHWPGRFGAALTGLGVNGDKLTDAIAGPEOENQGAQVYIF 555
 QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLILRSQ 600
 DB 556 YGASIASLSASHRIIGAHFSPGLQYFGQSLGGQDLTMDGLMDLAVAGQGLLLRAQ 615
 QY 601 PVLVRKAIMFNPREVARNVFECDQVVGKEAGEVRVCLVHVKSTRDRLEQIOISVVT 660
 DB 616 PVLRLATMEFSPKVARSVFACQEQVLKKNKAGEVRVCLVRKNTKDLREGDIQSTVT 675
 QY 661 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVRLNF 720
 DB 676 YDLALDPVRSRIRAFDETQNTNRRRTQVFLMKQKCELTLLIPDCVDSVSPILRLNY 735
 QY 721 SLVGTPLSAGNLRPVLAEDAQRLFTALPFKCNKGNNDICQDLSITFSFMSLDCLVWG 780
 DB 736 TLVGEPLRSFGLNLRPVLAEDAQRFETAMFPFKCNKGNDSICQDLSITFSAMGLDVLWG 795
 QY 781 GPREFNVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWL-ACESASSTE 839
 DB 796 GPQDFNMSVTLRNDGDSYGTQVTVYPSGLSYRKQASQNPITKPKFWKPAESSSSSE 855
 QY 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNTE 899
 DB 856 GHGALKSTTNINHPIPPANSEVTFNVTFDVDSHASPGLKLLKAI VASENNMSRTHKTK 915
 QY 900 FQLELPKYAVVWVTSYKVLNFTASENTRVMOHQYQVSNLQORSILPSLVLPVP 959
 DB 916 FQLELPKYAIYIMVTSDESSIRYLNFTASEMSTKVIHQYQVFNQLQORSLPSVFWPVP 975
 QY 960 VRLNQTVMRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNVCSIAVCORICDI 1019
 DB 976 VQINNVTWDPQVIFSONLSSACHTEQKSPPHSNFRDQLERFVLNCSVAVCKRIQCDL 1035
 QY 1020 PFFGIOEFENATLKNLSFDWYIKTSHNHLILVSTAEILFNDSVFTLLPQOGAFVRSOTE 1079
 DB 1036 PSFNTQEIFNVTLKNLSFDWYIKTSHGHLILVSTAEILFNDSAFALLPQOESYVRSKTE 1095
 QY 1080 TKVEPFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEQ 1137
 DB 1096 TKVEPYEVRNPVPLIVGSSIGGLVLLALITAGLYKLGFFKRYQKDMNEAAPDAPPO 1153

```
RESULT 3
ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11c.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95.";
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81695; AAA59180.1; -;
DR EMBL; Y00093; CAA68283.1; -;
DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AAA51620.1; ALT SEQ.
DR EMBL; M29482; AAA51620.1; JOINED.
DR EMBL; M29483; AAA51620.1; JOINED.
DR EMBL; M29484; AAA51620.1; JOINED.
DR EMBL; M29485; AAA51620.1; JOINED.
```

```
DR EMBL; M29486; AAA51620.1; JOINED.
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163 INTEGRIN ALPHA-X.
FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128 POTENTIAL.
FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 165 351 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 POTENTIAL.
FT CA_BIND 530 538 POTENTIAL.
FT CA_BIND 593 601 POTENTIAL.
FT SITE 1131 1135 GFPR MOTIF.
FT DISULFID 69 76 BY SIMILARITY.
FT DISULFID 108 126 BY SIMILARITY.
FT DISULFID 655 712 BY SIMILARITY.
FT DISULFID 771 777 BY SIMILARITY.
FT DISULFID 848 863 BY SIMILARITY.
FT DISULFID 998 1022 BY SIMILARITY.
FT DISULFID 1027 1032 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490 G -> A (IN REF. 2).
FT CONFLICT 756 756 L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;
```

Query Match 58.8%; Score 3459; DB 1; Length 1163;
Best Local Similarity 60.9%; Pred. No. 9.9e-229;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

```
Qy 1 FNLDTENAMTFEENARGGQSVVQVQGSRRVVVGAQEIIVAAVNAORGSLYQCDYSTGSCPEI 60
Db 20 FNLDTEELTAFFVADVSGSVVQVQGVVGVGAQPKITAAVQVQGLYQCGYSTGACPEI 79
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLGSLNLROQPOK 120
Db 80 GLQVPVEAVNMSLGLSLASTTSPSOLLACGPTVHHECGRNWYLTGLCFLLGPT--QLTOR 137
Qy 121 FPEALRGCPQEDSDIAFLDGSGSIPIPHDFRMEKMWSTVMEQLKKSKTLFSLMQYSEFP 180
Db 138 LPVSRQECPRQEQDIVFLDGGSGSISSRNFMATMNFVRAVISOQFORPSTQFSLMQFSNKF 197
```

181 RIHFTKEFQNNPRSLIKIPITOLLGTRHTATGLRKVVRLEFNITNGARKNAFKILFL 240
198 QTHFTPEFRRTNPLSLLASVHQLQGFYTTATAIQNVVRHLPHASVGARDATKILIVI 257
241 TDGEKFGDPIGYEDVPELDRGVIYRIVGVDAFRSEKSRQELNVTASKPRPDHVFQIN 300
258 TDGKKEGSDLDYKDVIPMAADAAGIIRVAIGVLAFAQNRNSWKLNDIASQSEHIFKVE 317
301 NFALKTIQNLREKIPAIETGQTGSSSEFEHSEMSQEGFAAITSNGPLLSSTVGSYDMAG 360
318 DFDALDQNLQEKIFAIETGTTSSSEFEHSEMSQEGFAVFTPDGPFVLAGVGSYFWSG 377
361 GVFLYTSKESKSTFNMTVRDSDMDNDAVLGVAAGAAIILNRVQSLVLAGAPRYOHLGLVAMFR 420
378 GAFLYPNMSPFTFNMDSQENVDNRDSYLGSTELALWKGVQSLVLAGAPRYOHTGVIF 437
421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQVSVCP 480
438 QVSRQWRKAEVTTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYYEQTGGQVSVCP 497
481 PRGORARWOCDAVLIGQCGQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAYVLF 540
498 PRGWR-RWCDVLYGSGHPWGRFGAALTVLGVNGDKLTDVVGAPGEEENRGAYVLF 556
541 HGTSGSISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLLLRSQ 600
557 HGVLPISPSHSORIASGLSSRLQYFGQALSGGQDLTDGLVDLAVGARGQVLLLR 616
601 PVLVKAIMPEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLVGVVSMQIPAEIPRSFAFCEQVQVSEQLTQVSNICLYIDKRSKNLLGSRDLQSSVT 676
661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCEIDPVPVILBNF 720
677 LDALDPCRLSPRATFOETKRSLSRVVLGKLKACNFENLLPSCVEDSVTPITLRLNF 736
721 SLVGTPLSAFNLPRVLAEDAOFLTALFPPEKNCNDNICODDLSITFFSMDCLVWG 780
737 TLVGKPLAFNLPRVLAEDAOFLTALFPPEKNCNDNICODDLSITFFSMDCLVWG 796
781 GPREFNVTVVRNDEGDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWRLACASSTEV 840
797 SNLELAEVVMWMDGDSYGTITTFSPAGLSYRYVAEGQKQGLRSLHLTCDSPVGV-- 854
841 SGALKSTSCSINHIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSENNMPTNKTEP 914
901 QLELPVKYAVVMVTSKYSTKYNFTAS-ENTSRVMQHOYOVNGLQSRSLPISLVLPV 959
915 QLELPVKYAVVMVTSKYSTKYNFTAS-ENTSRVMQHOYOVNGLQSRSLPISLVLPV 974
960 VRLNQTWIDRPQVTSSENLSSTCHTKERLPSHDSFLAELKAPVNCSTIACVQRIQCDI 1019
975 VELQEAVMWDEVSHPQNSLRCSSEKIAAPPDSFLAHIQKNPVLDCSIAGCLRCFCDV 1034
1020 PFGIOBEFNATLKNLSFDWIKYKTSNHLILIVSTAETLNDVSTLPGQAFVRSQTE 1079
1035 PSFVSQBELDFTLKNLSFGVVRQILQKKVSWVSAETFTDTSVYSQLPQGEAFMRAQTT 1094
1080 TKVPEFVNPPLVLGVSSVGLLLALITLTAALYKLGFFKQYKDMSE 1128
1095 TVLEKRVNPTPLVGVSSVGLLLALITLTAALYKLGFFKQYKDMSE 1143

RESULT 4

ITAD HUMAN STANDARD; PRT: 1162 AA.
AC Q1349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
ITGAD.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=961111956; PubMed=8777714;
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
Stauton D.E., Gallatin W.M.;
RA "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3."
RT J. Biol. Chem. 275:8959-8969(2000).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Sp1 and Sp3."
RT J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human beta 2-integrin alpha subunit."
RT Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W., Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "Alphabeta2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)."
RT J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A., Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1."
RL J. Immunol. 163:1984-1990(1999).
CC -1- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES FROM THE BLOOD.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLEENIC RED PULP MACROPHAGES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; U37028; AAB38547.1; -

DR EMBL; U40274; AAB60634.1; --
 DR EMBL; U40275; AAB60635.1; --
 DR EMBL; U40276; AAB60636.1; --
 DR EMBL; U40277; AAB60637.1; --
 DR EMBL; U40278; AAB60638.1; --
 DR EMBL; U40279; AAB60639.1; --
 DR EMBL; U40278; AAB60638.1; JOINED.
 DR EMBL; AF187881; AAF62875.1; --
 DR HSP; P11215; IABX.
 DR Genew; HGNC:6146; ITGAD.
 DR MIM; 602453; --
 DR GO; GO:0008305; C.integrin complex; TAS.
 DR GO; GO:0004895; P.cell adhesion receptor activity; TAS.
 DR GO; GO:0016337; P.cell-cell adhesion; NAS.
 DR GO; GO:0007160; P.cell-matrix adhesion; NAS.
 DR GO; GO:0006955; P.immune response; NAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWF_A; 1.
 DR InterPro; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium;
 KW Magnesium.
 FT SIGNAL 1 17
 FT CHAIN 18 1162
 FT DOMAIN 18 1100
 FT TRANSMEM 1101 1124
 FT DOMAIN 1125 1162
 FT REPEAT 32 85
 FT REPEAT 7 2
 FT DOMAIN 150 332
 FT REPEAT 350 400
 FT REPEAT 401 452
 FT REPEAT 454 516
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 465 473
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1127 1131
 FT DISULFID 67 74
 FT DISULFID 106 124
 FT DISULFID 655 710
 FT DISULFID 769 775
 FT DISULFID 846 861
 FT DISULFID 994 1018
 FT DISULFID 1023 1028
 FT CARBOHYD 59 59
 FT CARBOHYD 87 87
 FT CARBOHYD 99 99
 FT CARBOHYD 391 391
 FT CARBOHYD 691 691
 FT CARBOHYD 733 733
 FT CARBOHYD 873 873
 FT CARBOHYD 957 957
 FT CARBOHYD 1046 1046
 FT CONFLICT 500 500
 FT CONFLICT 515 518
 FT CONFLICT 825 825
 FT CONFLICT 984 984
 FT SEQUENCE 1162 AA; 126885 MW; F296A1A3545D77D CRC64;

Query Match 57.8%; Score 3401; DB 1; Length 1162;
 Best Local Similarity 59.3%; Pred. No. 9.3e-225;

Matches :669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVLQGSRRVVVGPQEIIVAAQNGSLVQCDSYSGSCEPI 60

Db 18 FNLDVEEPTIFQEDAGGFGQSVVVGAPLEVVAAQNTQRLVDCAAATGMCQPI 77
 QY 61 RLQVEVEAVNNMGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFNSLRQQOK 120
 Db 78 PLHIREAVNNMGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFNSLRQQOK 136
 QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQKKSKTKLFSMLSEB 180
 Db 137 VPDATPECPHQMEDIIVFLIDGSGSIDQNDFNQMGFVQAVMGQFEGTDTLFAIMOYSNLL 196
 QY 181 RIHPTFKFQNNPNRSLIKPTTOLGRTHATGLRVKVRLEFNITNGARKNAPKILFLL 240
 Db 197 KIHFPTFOTRYSPOQSLVDPLVQLKGLTFTATGLTVTVTQLFHHKNGARKSAKILVI 256
 QY 241 TDGEKFGDPLGYEDVIPBLDREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVFOIN 300
 Db 257 TDGQKYDPLEYSDVIPOAEKAGIIRYVIGVGHAFQGTARQELNTISSAPQDHFVKVD 316
 QY 301 NFEALKTIQNLREKIFAIETGTQSSSPHEMSQEGFSAAITNSGPLLSTVSGYDWAG 360
 Db 317 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLPLGAVGFSMSG 376
 QY 361 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAIIILNRVQSLVLGAPRYOHILVAMPR 420
 Db 377 GAFLYPNNMSPPTFINMSQENVMDRSDYLGYSTELALWKGVQNLVGLAPRYOHTGKAVIT 436
 QY 421 QNTGHWESNANVKGTOIGAYFCASLCSDVDNSGSTDLVIGAPHYETRGQGVSVCP 480
 Db 437 QVSRQWRKKAQVGTQIGSYFCASLCSDVDNSGSTDLVIGAPHYETRGQGVSVCP 496
 QY 481 PRGQARWQCDVLYGEGQPMGRFGAALTVLGVDNGDKLTDVAIGAPCEEDNRGAVYLF 540
 Db 497 PRGQVQWQCDVLYGEGQPMGRFGAALTVLGVDNEDKLDVAIGAPCEQENRGAVYLF 556
 QY 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGSLSGQDLTMDGLVDLTVGAQHVLRLRSQ 600
 Db 557 HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLRSL 616
 QY 601 PVLRVKAIEMFNPREVARNFECNDQVYVKGAEGRVCLHVQKSTRRLREGQIQSVVT 660
 Db 617 PVLKGVAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLOQL--GDIQSSVR 674
 QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
 Db 675 FDLALDPGLTSRAIFNETKPTLTKRTKGLGHCETLKLPLDCVDDVVSPIILHLNF 734
 QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEPTALPPEKNGCNDNI CODDLSITFSFMSLCLVVG 780
 Db 735 SLVREPIPSQNLRPVLAEGSQDLPTASLPPEKNGCQDGLCEGLGLVTLFSGLQTLTVG 794
 QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
 Db 795 SSLEUNLVTVVWNADEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 853
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPNTRKTFEF 900
 Db 854 EG-LRSSRCSVNHPIFHSNGCTFIVTVDVSKATLGDRLMRASSENKASSKATF 912
 QY 901 QLELPVKYAVVMVTSVSHGVSTKYLNLF-TASNTSRVMQHVQVSNLQGRSLPISLVFLVP 959
 Db 913 QLELPVKYAVVMVTSVSHGVSTKYLNLF-TASNTSRVMQHVQVSNLQGRSLPISLVFLVP 972
 QY 960 VRLNQTVDWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQIOCDI 1019
 Db 973 VLLNGVAVDVMMEAPSOQL--PCVSEKRPQHSDFLTQISRSPMLDSCSIADCLQRCVDV 1030
 QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
 Db 1031 PSFSVQEBELDTLKNLSFGVWRETLOKKVLVSVVAEITFTDSVYSQLPQGEAFRAQWE 1090
 QY 1080 TKVEPFEVNPBLPLIVSGSVGLLLALLIITAAALYKLGFFKQYKDMMSSE 1128

Db 1091 MULEDEYVNAIPINGSSVGALLLALITATLYKLGFPRHYKEMLED 1139

RESULT 5

ITAL_HUMAN STANDARD; PRT; 1170 AA.

AC P20701; O43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1

DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)

DE (CD11a).

GN ITGAL OR CD11a.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=2537322;

RA Larson R.S., Corbi A.L., Berman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1

RT alpha subunit: an integrin with an embedded domain defining a protein

RT superfamily.";

RL J. Cell Biol. 108:703-712(1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=94425270; PubMed=10493829;

RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,

RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,

RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,

RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from

RT human chromosome 16p and 16q";

RL Genomics 60:295-308(1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Ou A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha

RT L beta 2) integrin.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398682; PubMed=8805579;

RA Ou A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain

RT from the CD11a/CD18 integrin.";

RL Structure 4:931-942(1996).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=94425288; PubMed=10493852;

RA Kallen J., Weizenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,

RA Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the

RT CD11a'I-domain'.";

RL J. Mol. Biol. 232:1-9(1999).

CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,

CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA

CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL

CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES

CC AND MONOCYTES.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L

CC ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=P20701-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P20701-2; Sequence=VSP_002738;

CC Note=No experimental confirmation available;

CC

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide cdlla.htm".

CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cdlla.htm".

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC

CC EMBL: Y00796; CAA68747.1; -

CC EMBL: AC002310; AAC31672.1; -

CC PIR: S03308; S03308.

CC PDB: 1LFA; 29-JAN-96.

CC PDB: 1ZON; 07-DEC-96.

CC PDB: 1ZOO; 07-DEC-96.

CC PDB: 1ZOP; 07-DEC-96.

CC PDB: 1CQP; 07-AUG-00.

CC PDB: 1DQO; 03-FEB-00.

CC PDB: 1MJN; 28-JAN-03.

CC PDB: 1MO9; 14-JAN-03.

CC PDB: 1MQA; 14-JAN-03.

CC Genew: HGNC:6148; ITGAL.

CC MIM: 153370; -

CC GO: GO:0008305; C:integrin complex; TAS.

CC GO: GO:0006928; P:cell motility; TAS.

CC InterPro: IPR000413; Integrin_alpha.

CC InterPro: IPR002035; VWF_A.

CC Pfam: PF01839; FG-GAP; 3.

CC Pfam: PF00357; Integrin_A; 1.

CC PRINTS: PRO0092; VWF; 1.

CC PRINTS: PRO1185; INTEGRINA.

CC PRINTS: PRO0453; VWFADOMAIN.

CC SMART: SM00191; Int_alpha; 4.

CC SMART: SM00327; VWFA; 1.

CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.

CC PROSITE: PS0234; VWFA; 1.

CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

CC Signal; 3D-structure; Magnesium; Calcium; Repeat;

CC Alternative splicing.

CC SIGNAL 1 25

CC CHAIN 26 1170 INTEGRIN ALPHA-L.

CC DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 1089 1112 POTENTIAL.

CC DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

CC REPEAT 42 91 FG-GAP 1.

CC REPEAT 92 149 FG-GAP 2.

CC DOMAIN 170 349 VWFA.

CC REPEAT ? ? FG-GAP 3.

CC REPEAT 401 455 FG-GAP 4.

CC REPEAT 457 516 FG-GAP 5.

CC REPEAT 518 575 FG-GAP 6.

CC REPEAT 578 630 FG-GAP 7.

CC CA BIND 468 476 POTENTIAL.

CC CA BIND 530 538 POTENTIAL.

CC CA BIND 590 598 POTENTIAL.

CC SITE 1115 1119 GEFKR MOTIF.

CC DISULFID 73 80 BY SIMILARITY.

CC DISULFID 111 129 BY SIMILARITY.

CC DISULFID 653 707 BY SIMILARITY.

CC DISULFID 771 777 BY SIMILARITY.

CC DISULFID 845 861 BY SIMILARITY.

CC DISULFID 998 1013 BY SIMILARITY.

CC DISULFID 1021 1052 BY SIMILARITY.

CC CARBOHYD 65 65 N-LINKED (GLCNAC...) (POTENTIAL).

```

FT CARBOHYD      89      89      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     188     188      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     649     649      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     670     670      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     726     726      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     730     730      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     862     862      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     885     885      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     897     897      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD    1060    1060      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD    1071    1071      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC     954     954      O -> QGVHGLVEMOTSKOILCRPAGDAEHTVGAQEGELPC
FT                                     /FTIGVSP 002738.
FT                                     R -> W (IN REF. 1 AND 2) .
FT                                     Y -> I (IN REF. 2) .

FT CONFLICT      214     214
FT CONFLICT      660     660
FT STRAND        155     155
FT STRAND        164     164
FT STRAND        166     166
FT TURN          169     169
FT TURN          185     185
FT TURN          188     188
FT STRAND        191     191
FT STRAND        202     202
FT STRAND        208     208
FT STRAND        213     213
FT TURN          217     217
FT TURN          222     222
FT STRAND        229     229
FT STRAND        233     233
FT TURN          243     243
FT TURN          244     244
FT TURN          249     249
FT TURN          250     250
FT TURN          253     253
FT STRAND        256     256
FT TURN          274     274
FT TURN          277     277
FT STRAND        280     280
FT TURN          288     288
FT TURN          293     293
FT TURN          298     298
FT TURN          300     300
FT TURN          302     302
FT TURN          310     310
FT TURN          311     311
FT TURN          318     318
FT TURN          319     319
FT TURN          328     328
FT TURN          329     329
FT TURN          330     330
SQ SEQUENCE     1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match      26.3%; Score 1547.5; DB 1; Length 1170;
Best Local Similarity 34.3%; Pred. No. 9.le-98;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

QY      1  FNLDTENAMTFO--ENARFGOSVVOLOGSRVVVGAPQBIVAANORGSLYQCDYSTGSCE 58
DB      26  YNLDRVGARSPFPGRAGRIFGRVLQV--GNGVIVGAPGE---GNSTGSLYQCSGTGHCL 81
QY      59  PIRLOVPEAVNMSLGLSLAATTPSPQLLACGPTVHQTCSENTYKGLCFGLFGSNLR--- 115
DB      82  PVTLR--GSNYTSKYLGTWLTATDPTDGSILACDPLGSRCTCDQNTYLSGLCYLFRQNLQGP 140
QY     116  -QOPQFPALRGCPDESDIAFLIDGSGSIITHDFRRMKWVSTVMEOLKSKSLFSLM 174
DB     141  LQGRPGFORCIRG---NVDLVFLFDGSMLOPDEFOKILDFMKDVKKLSNTSYQFAAV 196
QY     175  QYSEEPRIHPTKPEFONNPNRSLIKPITQLGRHTATGLRVVRELENTINGARKNAF 234
DB     197  QSTSTKTFDESDYVKRQDPDALLKHVHMLLTNTFGAINVATEVFEELGARPDAT 256
QY     235  KILFLITDGEKFGDPLGYEDVIPELDREGVIRVIGVGDAFRSEKSRQBLNTVASKPPRD 294
DB     257  KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIKGHFQTKESQETLHKFASKPASE 309
QY     295  HVFQINNFBALKTIQNLREKIFAIBGTQTGSSSPFHEMSQEGFSAATISNGPLSTVSG 354

```

```

310 FVKILDTPEKLDLFTLEQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVVGAVG 369
355 SYDWAGGVF--LYTSKEKSTFINMTRVSDMDMDAYLGAAA--IILNRVQSLVGLGAPRYQH 412
370 AKDWAGGFLDLKADLQDDTFIFGNEPLTPEVRAGLYGTVTWLPSRQKTSLLASGAPRYQH 429
413 IGLVAMFR--QNTGWESNANVKTGICAYFGASLCSDVDVDSNGSTDVLVIGAPHYHQT 470
430 MGRVLLFQEPQGGHWSQVQIHTGTQISYFGGELCGVDVDQDGETELLIGAPLYGEQ 489
471 RGGQVSVCPFRGQRARWQCDV--LYGEOQPMRFGAALTVLGDVNGDKLTDVAIGAP 528
490 RGRVFIY----QRRQLGFEVSELOQDGPVPLGRFGEAITALTDINGDLVDVAVGAP 544
529 GEEDNRGAVYLFHGTSGSGISPSHQSRIAGSKSLPRLQYFGQSLSGGQDLTMDGLVLDTV 588
545 LEE--QGAVYIFNGRHG--GLSPQSPQRIEGTQVLSGIQWFGRSIHGVNDLEGDGLADVAV 601
589 GAQGHVLLRSQPVLRVKAIMEFNPREVARNVFCNDQV--KGKAGEVRVCLHVQKSTR 647
602 GAESQMIVLSSRPVDMVTLSFSPAETPVHVEVECSYSTSNMKKEGVNITICFQI--KSLY 660
648 DLRREGQIQSVVTVYDLALDSGRPHSRAVFNETKSTRQTQVLGTQTCTELKQLPNCI 707
661 PQF--QGRIVANLTYTLQDGHTRRRGLFPGGRHELNRNIAT--TSMCTDFSFPVVCV 718
708 EDPVSPVILRLNFSL---VGTPLS--AFGN-----LRPVLAEADAQRLFTALFPFKNCGN 757
719 QDLISPINVSLNFSLWEEBEGTPDQRAQOKDIPPIRLPSLHSETWEI-----PFEKNCGE 773
758 DNICQDLSITFSFMSLDCVLVGGPREFNVTVVNRDGEDSYRTQVTFPPFLDLRYKVS 817
774 DKCEANLRVSFSPARSALRLTAPASLSBELSLNLEEDAYVWQLDHFPPGLSFRKVE 833
818 TLQORSORSWRLACES--ASSTEVSGALKSTCSINHPIPENSEVTFNITFDVDSKAS 875
834 ML---KPHSQIPVSCCELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887
876 LGNKLKLLKANVTSENN---MPRTNKTEFQLELPVKYAVVWVTVSHGVSTKYLNFTASEN 931
888 WGDVSELHANVTCNNEEDSDLEDNSATTI---IPILYPINILIQDESDTLVYSTPKGP 944
932 TSRVNHQHOYQV---SNLQGRSLP--ISLVFLVPLVRLNQTVIWDPRPOVTFSENLSSTCHTK- 986
945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPPQPPSEGPITHOMSVQMEPPV--PCHVED 1002
987 --ERLPESHSD--FLAELRKAPVUNCIAVCQRIQCDIPFGIQEENFATLKNLSFDWYIK 1043
1003 LERLPDAAEPCLPGALFPCPV-----FRQEILVQVIGTLELVGEIE 1044
1044 TSHNHLIVSTAELIFNDSVFTLLPGOGAFVRSQTETKVEPPEVPNPPLIIVGSSVGGLL 1103
1045 AS-SMFSLCSSLSISFNSKHFHLYGSNASL-AQVMKVVDVYKQMLYLYLVLSIGIGLL 1102
1104 LLALITAAALYKGLFPKRYKDMMSSEG-QPPGAEP 1136
1103 LLLLIFIVLYKVGFPKRLKEKMEAGRVGPNIP 1136

RESULT 6
ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=91268576; PubMed=2051027;
Kaufmann Y., Tseng E., Springer T.A.,
"Cloning of the murine lymphocyte function-associated molecule-1
alpha-subunit and its expression in COS cells.";
J. Immunol. 147:369-374(1991).
[2]
SEQUENCE OF 24-42.
MEDLINE=9518276; PubMed=3887182;
Springer T.A., Teplow D.B., Dreyer W.J.,
"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
glycoproteins and unexpected relation to leukocyte interferon.";
Nature 314:540-542(1985).
-!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
LEUKOCYTES RECRUITMENT.
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
ASSOCIATES WITH BETA-2.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- TISSUE SPECIFICITY: LEUKOCYTES.
-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
-!- SIMILARITY: Contains 1 VFMA domain.
-!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M60778; AAA39426.1; -
PIR; I56126; I56126.
HSSP; P20701; 1LFA.
MGD; MG1:96606; Itgal.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWF_A.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00357; Integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS50234; VWF_A; 1.
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Magnesium; Calcium;
Repeat.
SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT CA_BIND 528 536
POTENTIAL.
INTEGRIN ALPHA-L.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
FG-GAP 1.
FG-GAP 2.
VWFA.
FG-GAP 3.
FG-GAP 4.
FG-GAP 5.
FG-GAP 6.
FG-GAP 7.
POTENTIAL.
POTENTIAL.

FT	CA_BIND	588	596	POTENTIAL.
FT SITE	1111	1115		GFPR MOTIF.
FT DISULFID	70	77		BY SIMILARITY.
FT DISULFID	108	126		BY SIMILARITY.
FT DISULFID	147	199		BY SIMILARITY.
FT DISULFID	651	705		BY SIMILARITY.
FT DISULFID	767	773		BY SIMILARITY.
FT DISULFID	840	856		BY SIMILARITY.
FT DISULFID	993	1009		BY SIMILARITY.
FT DISULFID	1017	1048		BY SIMILARITY.
FT CARBOHYD	86	86		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	185	185		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	270	270		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	444	444		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	668	668		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	696	696		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	724	724		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	728	728		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	776	776		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	857	857		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	880	880		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	890	890		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	899	899		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	927	927		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1056	1056		N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE	1163 AA;	128343 MW;	A7A3078489E8232F CRC64;	

Query Match 26.0%; Score 1527.5; DB 1; Length 1163;
Best Local Similarity 34.1%; Pred. No. 2.1e-96;
Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;

Qy	1	FNLDENAMTFOENA-RGFQSVVQLQGSRRVVVGAPEIIVAAHQRLSLYQCDYSTGCEP	59
Db	24	YNLDTRPTQSLAQAGRHFGYQLQIEDG-VVVGAPGE--GDNTGLYHCRTSSEFCQP	79
Qy	60	IRLQVPEAVNMLSLAATTPPOLLAGCTVHTCSENTYVKGCLFGLFSGNLKQQPQ	119
Db	80	VSLH-GSNHTSKYLGMTLATDAAKGSLACDPLSRTCDQNTYLSGLYLFPOSLGPM	138
Qy	120	KPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEMVSTVMEQLKSKTFLSLMOYSEE	179
Db	139	QNRPAVOECMKGVLDVFLFDGQSLDRKDQFELBFMDKVMKKNSTSYQFAAVQFSTD	198
Qy	180	FRHFTFKFQ-QNNPNRSLIKPIETOLLGRTHATGLRKYVRLFNITNGARKNAKILF	238
Db	199	CRTEFTFLDVVKQKPNVLLGSVQPMFLLTNTFRATVVAHVFEESGARPDATKVLV	258
Qy	239	LLTDGKFP--GDPLGYEDVIPLEDRGVTRYVIGVDGAFRSEKSRQELNTVASKPRDHV	296
Db	259	IITDGEASOKGNISAAHD-----ITRYIIGIKGHFVSVQKQKTLHIFASEPVEEFV	309
Qy	297	FQINNEALKTIQNLREKIFATEGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSY	356
Db	310	KILDTFEKLKDLFTDLORRIAYIEGTNRDLTSFNMLSSGSLADLSKGHAVVGAVGAK	369
Qy	357	DMAGGVF-LVTSKSEKSTFNMTVDSDMDNDAYLGAYAAA-IILNRVQSLVGLGAPRYOHIG	414
Db	370	DMAGGFLDREDLOGATFVQGEPLTSDVRGGYLGTVVAMWTSRSPRLAAGAPRYOHVG	429
Qy	415	LVAMFR--QNTGMWESNANVKGQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYYQTRG	472
Db	430	QVLLFQAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGAELLIGAPLFFGEQRG	489
Qy	473	GVSVCPPLPRGQARWQCDVLYGEOGPWGRGAALTVDVGVNGDKLTDVAIGAPGEED	532
Db	490	GRVFTY---QRRQSLFEMVSELQGDGPGYPLGRFGAAITALTIDNGDLTDVAIGAPLEE-	545
Qy	533	NRGAVYLFHCTSGSGISPHSHQRIAGSKLSPRQYFGQSLSGQDLTMDGLVDLTVGAQG	592
Db	546	-QAVYIFNKGPG-GLSPQSPQRIQGAQVFPGRFWRFGRSIHGVKDLGGDLADVVVGAEG	603
Qy	593	HVLLRSQPVLRVKAINMEFNPREVARNVFCNDQVWKGKBAQ-EVRVCLHVQKSTRDLR	651

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;
 Query Match 19.5%; Score 1148.5; DB 1; Length 1167;
 Best Local Similarity 28.8%; Pred. No. 2e-70;
 Matches 355; Conservative 214; Mismatches 459; Indels 205; Gaps 43;
 QY 1 FNLDTEA--MTFENARGFQSVVQLQGGVVVGAQPEIVAAHQGS-----LYOCDY 52
 DB 20 FNMVDVWATLQGPAPAVLSSLLHLDPSN-----NQTCLLVARRSSNRTAALYCAI 74
 QY 53 STGCEPIRLQVPVEAVNMSLGLSLAATT--SPQLLAC-GPTVHQCSTYVYKGLCFL 109
 DB 75 SI-SPDEIACQ-PVEHICMPKRGYQGVTLVGNHNGVLVCVQVQARKFSLNSELGTACSL 132
 QY 110 FGSNLRQOPKPPALRG-----C-----PQE 131
 DB 133 LTPNLDLQAQAYFSDLEGFLDPGAHVDSGDYCRSGKSGTGEEKSARRRRRTVEEED 192
 QY 132 DSDIAFLIDGSGSIIPHDFRMKEWSTVMEQL--KSKTFLSLMOYSEFRHFTPEKF 189
 DB 193 GTEIAIYLDGSGSLGSDPQAKNPISTMMNFEKFCFECNFALVQVAVIOTFEDLOES 252
 QY 190 QNNPNRSLIKPIITQLLGRTHATGLRKVRELFNITNGARKNAKILFLITDCEKFGDP 249
 DB 253 RDINASLAKVQSIQVKEVTKTASAMQHVLDNIFIPSRGSRKKALKVMVLTGDIIFGDP 312
 QY 250 LGYEDVPELDRGVIRVIGVADAPRSEKSRQELNVTASKPRDRHVFOINNFPALKTIO 309
 DB 313 LNLTVINSRPMQGVAFVAFVIGVDRFKNNTYRELKLIASDPKEAHTFKVTNYSALDGLL 372
 QY 310 NOLREKFAIGTQSGSSSEHEMSOEGSAAITSNGP-LLSTVGSYDVGAGVFLY-TS 367
 DB 373 SKLQOIRVHMEGT---VCDALQYQLAOTGSAQLDKGOVLGTVGAFNWSGGALLYSTQ 429
 QY 368 KEKSTFINMT-RVDS-DKNDAYLGYAAAIILNRNVQSLVGLGAPRYQIHLVAMFRONTGM 425
 DB 430 NGRGCLNQTAKEDSRVQVSYLGYSLAVLHKAHGISVAGAPRHLRGAVFELRKEDR- 488
 QY 426 WESNA---NVKGTQIAYFGASLCSVDVDSNGSTDLVLIGAPHYVYETRGQGVSCPLPR 482
 DB 489 -EEDAFVRRIEGEMGVSFGVLCFVDIDMDGTTDFLLVAAPFYHIRGEGRVYVQVPE 547
 QY 483 GQARWQCDVLYGEGQGWCFGAALTVLGDVNGDKLTDAVAGP-----GREDNRGA 536
 DB 548 -QDASFSLAHTLSGHPGLTNRFGFAMAAGDINQDKFTDVAIGAPLEGFAGGAGSAGS 606
 QY 537 VYLFHGTSGGISPSHSORIASGKSLSPLOYFGOSLGGQDLTMDGLVLDLTVGAQGHVLL 596
 DB 607 VYIYNGHSG-GLYDPSQOIRASSVASGLHYFGMSVSGGLDFNGDGLADIIVGSRDPAV 665
 QY 597 LRSQPVLRVKAIMEFNPREVARNPECNQVVKKEAGEVRLVHVKQS---TRDLREG 653
 DB 666 LRSRVDLTVSMFTFP-----DALPMVFIGKM--DVNLCEVDSSVVASBFLGLEM 715
 QY 654 QIQSVTVYDLALDSRPHSRVAFNETKNSRTRQVGLGLTQC----- 696
 DB 716 FLNFTVDV-----TKORQLQCESSGCSCLRNKNGSFLCEHFWLI 760
 QY 697 ETLQLPNCIEDPVSPIVRLNLSVGLTPLSAFONLR-----PVLAEQAQLFTALF--P 750
 DB 761 STEEL-----CEDCFSNITIKVYE-----FQTSGRDRDYNPTL--DHYKEPSAIFOLP 809
 QY 751 FEKNCNDNI CODDLISITFSRMSLDCLVVGGPRFNVTVVRNDEGSYRQVTFPPFLD 810
 DB 810 YEKDCKNVFCIAEIQLTN-ISOQELVVGVTKEVTWNISLTNSGDSYMTNMAALNPRN 868
 QY 811 LSYKVTQLQVRSRNLACESASSTEVSCALKSTSCSNHPIFFENSEVTNITFDV 870
 DB 869 LQFKKI-----QKVPSPDQDDPKPV---ASVLVMNCKIGHPL-LKRSSVNSVVTWQL 918
 QY 871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVVWVVTSHGVSTKYNFTASE 930

RESULT 8

ITAE_HUMAN STANDARD; PRT; 1179 AA.
 ID ITAE_HUMAN STANDARD; PRT; 1179 AA.
 AC P38570; Q9NZU9;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-8 precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
 DE antigen (CD103 antigen) (Integrin alpha-IEL).
 GN ITGAE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
 RC TISSUE=Leukemia, and Lymphocytes;
 RX MEDLINE=941164962; PubMed=8119947;
 RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
 Parker C.M.;
 RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
 RL J. Biol. Chem. 269:6016-6025 (1994).
 RN [2]
 RP REVISIONS TO 88-114.
 RA Parker C.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 53-1179 FROM N.A.
 RC TISSUE=Fetal kidney;
 RX MEDLINE=20138496; PubMed=10673275;
 RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
 Stotlersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
 Green E.D.;
 RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
 RL Genome Res. 10:165-173 (2000).
 RN [4]
 RP MUTAGENESIS OF ASP-109 AND PHE-316.
 RX MEDLINE=20400502; PubMed=10837471;
 RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
 Brenner M.B.;
 RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
 RL J. Biol. Chem. 275:25652-25664 (2000).
 CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROM; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L25851; AAB59359.2; --
 CC DR EMBL; AF168787; AAF43107.1; --
 CC DR PIR; A53213; A53213.
 CC DR HSSP; P11215; IABX.
 CC DR Genew; HGNC:6147; ITGAE.
 CC DR MIM; 604682; --
 CC DR GO; GO:0008305; C:integrin complex; TAS.
 CC DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC DR InterPro; IPR000413; Integrin_alpha.
 CC DR InterPro; IPR002035; VWF_A.
 CC DR Pfam; PF01839; FG-GAP; 3.
 CC DR Pfam; PF00357; Integrin_A; 1.
 CC DR Pfam; PF00092; vwa; 1.
 CC DR PRINTS; PR01185; INTEGRINA.
 CC DR PRINTS; PR00453; VWFADOMAIN.
 CC DR SMART; SM00191; Int_alpha; 3.
 CC DR SMART; SM00327; VWA; 1.
 CC DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC DR PROSITE; PS0234; VWF_A; 1.
 CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Polymorphism; Magnesium; Calcium.
 FT SIGNAL 1 18
 FT CHAIN 19 1179
 FT CHAIN 19 1179
 FT CHAIN 19 1177
 FT CHAIN 179 1177
 FT CHAIN 179 1124
 FT DOMAIN 1125 1147
 FT TRANSMEM 1148 1179
 FT DOMAIN 181 198
 FT REPEAT 181 198
 FT REPEAT 2 2
 FT REPEAT 2 2
 FT DOMAIN 145 199
 FT DOMAIN 200 391
 FT REPEAT 401 456
 FT REPEAT 457 506
 FT REPEAT 510 571
 FT REPEAT 573 638
 FT REPEAT 641 693
 FT REPEAT 641 693
 FT CA_BIND 522 530
 FT CA_BIND 586 594
 FT CA_BIND 654 662
 FT SITE 1150 1154
 FT SITE 70 79
 FT DISULFID 126 159
 FT DISULFID 706 762
 FT DISULFID 823 829
 FT DISULFID 893 907
 FT DISULFID 1008 1033
 FT DISULFID 1041 1057
 FT CARBOHYD 49 49
 FT CARBOHYD 271 271
 FT CARBOHYD 321 321
 FT CARBOHYD 444 444
 FT CARBOHYD 726 726

FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 360 360 D -> E.
 FT VARIANT 1041 1041 /FTID=VAR_008884.
 FT VARIANT 1041 1041 C -> S.
 FT MUTAGEN 208 208 /FTID=VAR_008885.
 FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.
 FT MUTAGEN 477 477 F->A: LOSS OF E-CADHERIN BINDING.
 FT CONFLICT 482 482 V -> I (IN REF. 3).
 FT CONFLICT 482 482 Q -> R (IN REF. 3).
 FT CONFLICT 950 950 R -> W (IN REF. 3).
 FT CONFLICT 1019 1019 R -> V (IN REF. 3).
 SQ SEQUENCE 1179 AA; 13008 MW; E558902EDF9D95E1 CRC64;
 Query Match 19.4%; Score 1140; DB 1; Length 1179;
 Best Local Similarity 28.9%; Pred. No. 7, 7e-70;
 Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;
 QY 45 GSLYQCOYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTPPQLLACGPTVHQ 95
 DB 65 GPLHRCSLVQDEILCHPVEHVPIKGRHGVTVVRSHHGVLICI-----QVLVRRP--HS 117
 QY 96 TCSENYVKGLCFLEGSNLRQPO-----
 DB 118 LSSELT---GTCSLLGPDRLRPOAQNFFDLENLLDPDARVDVTDGDCYNKEGGEDDVNTA 174
 QY 120 KPPEALRGCPQED-----SDIAPLDGSGSIIPHDFFRMKEWSTVMBQL-- 164
 DB 175 RRRALKEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEED 234
 QY 165 KSKTLFSLMOYSEEFRIHFTKFPQNPNSLIKPTQLLGRTHRTHTGLRKYVRELFN 224
 DB 235 KCFCNFALVOYGVQVTEFDLRDSQDVMSLARVQNIQVGSVTKTASAMQHVLDSIFT 294
 QY 225 ITNGARKNAFKILFLLTDEKEGDPGLGYEDVPELDREGVIRYVGVGDAFSEKSRQEL 284
 DB 295 SSGSRKASKVMVVLTDGGIFEDPLNLTIVNSPMQGVFERFAIGVGEFESKARTAREL 354
 QY 285 NTVASKPRDRHVQINNFALKTIQNLREKIFAETGTQTSSTSSSEFHEHMSOEGFSAAT 344
 DB 355 NLIASDPETHAFKVTNYWALDGLSKLRNYIISMEGT---VGDALHYQLAQIGFSAQL 411
 QY 345 SNGP--LLSTVGSYDWAGGVFLY--TSKEKSTFTINMTRVDSMDNDA----YLGVAAILRN 398
 DB 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAADAAADAAQAQSYLGVAVAVLHKT 471
 QY 399 RVQSLVLCAPRYQHIGLVAMFR--QNTGMWESNANV--KGTQICAYFGASLCSDVDVDSNGST 456
 DB 472 CSLSYVAGAPQYKHHG--AVFELQKEGREAFLPVLEGEQMGSYFGSELCPVDIDMDGST 529
 QY 457 DLVLIGAPHYBQTRGGOVSCVPLPRGORARWOCDAVLGEOGQPMGRFGAALTVLGDVN 516
 DB 530 DFLVLAAPFYVHGEGRVYVYRLSE--QDGSFLARILSGHGFNFTNAREFGFMAANGDUS 588
 QY 517 GDLKTDVAIGAP---GEEDNR--GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQ 570
 DB 589 QDKLTDVAIGAPLEGFGADDGASFGSVIYNG--HWDGLSASFSQIRASTVAPGLQYFCM 647
 QY 571 SLUGGQDLTMDGLVDLTGVAQGHVLLLRQPVLRVKAIMEFNPREVARNVFCNQDVVKG 630
 DB 648 SMAGGFDISGDLADITVGTLLQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
 QY 631 KEAGSVRVLCHVQKSTRDRREGQIQSVVTVYDLADSGRPHSRVAFNETKSTRQTQVL 690
 DB 698 NGVNVRLCFEI--SVTTTASEGLREALNFTLDVVDVGKRRRLQCSVDVRSCLGCLREWS 756
 QY 691 GLUTQCEITKLQLPN---CIEDPVSPIVLRNFSLVGTPLSAFNGRLRPLVLAEDAQRLFT 746
 DB 757 SSGQLCEDL-LLMPTGEGLCEDECDFSNASVKVSYQL--QTPEQTDHPQDILDRYTPPAI 814

QY 107 CFLPGNLRQOPKFPALRGCCQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTWBLQKK 166
DB 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVILVDGNSIYPWDS-----VTAFLNLDLX 170
QY 167 -----SKTLFSLMOWSEEFRIHFTFKEQNPNPRLIKPIITOLLGR--THRTATGLRKVV 219
DB 171 RMDIGPKQTVQVIVQGENWTHFNLNKYSSTSEVLVAAKKIVORGGRTMTALGDTAT 230
QY 220 RELFNITNGARKNAFKILFLLTDGKFGDPLGVEDVPELDREGVIRYVIGVDAPR--- 276
DB 231 KEAFTEARGARRGVKKVMVITVDGESH-DNHLRLKKVIQDCEDENIQRFSAIILGSYNRGN 289
QY 277 --SEKSRQELNVTASPPDRHVPQINNFEALKTIONLREKIPAIETGOTGSSSPEHEM 334
DB 290 LSTEKEVEEIKSTASEPTEKHPFNVSDELALVTIVKTLGERIPALBATAQDAAAPSEM 349
QY 335 SQEGFSAATTSNGPLLSLVGSDWAGGVFLYTSKE-----KSTF-INMTRVDSDNMDAYL 388
DB 350 SQTGFSAHYSQDWMLGAVGAYDWNGTVVQKASQIIIPRNTTFNVESKNEPL-ASYL 408
QY 389 GYAAAIILNRVQSL-VLGAAPRYOHIGLVAMFRQNTGMNESNANVKGTOIGAYFGASLCS 447
DB 409 GYTVNSATSSGDVLYIAGOPRYNHTGQVIYRMEDGNIKILQTLGSEIGSYFGSILIT 468
QY 448 VDVDSNGSTDLVLIGAPHY-----YEQTR-GGOVSVCPPLPRGORARWQCD 491
DB 469 TDIDKDSNTDILLVGAAPMTYGTKEKRGQKVYVVALNQTRFEVQMSLEPIKQTCSSRQHN 528
QY 492 AVLYGQGPWG-RFGAALTVLGDVNGDKLTDVAICAPGEONRGAVLPHGTSGSISP 550
DB 529 SCTTENKNEPCGARGFTAAVAKDLNDGFDNDIVIGAPLEDHGGGAVIYHG-SGKTIRK 587
QY 551 SHSQRITAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTGVAQGHVLLLRQPVLRVKAIM 610
DB 588 EYAQRIPSGDGKTLKFFQGSINGENDLNGDGLTDTIGLGAALFWRSDVAVVKVTWN 647
QY 611 FNPPEVARNVFECNDQVWKGKAG--EVRVCLHVQ-KSTRDLRREGQIOSVVTYDLALDS 667
DB 648 FEPNKVNIQKNCH--MEGKETVCINATVCFEVKLKSKEDTIYEADLQ----YRVTLDS 700
QY 668 GRPHSAVENET-----KNSTRQOVGLGTCTETLKLQLPNCIEDVPSIVLRLNFS 721
DB 701 LRQISRSFSGTQERKQVRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDPN 752
QY 722 LVGTPLSAFNLRLPVLAEADQRLFTALFPFEKNCNDNICQDDLSITTFMSLDCLVWGG 781
DB 753 LT-DPENG-----PVLDDSLPNSVHEYIIPAKDCGKKEKISDLSLHVATTEKDLIVRS 806
QY 782 PRE-FNVTVTVRNDGSDSYRTQVTFPPPLDLSYRKVSTIQNQRSQSRWSRLACESSTEV 840
DB 807 QNDKFNVSILTAKNDSAYNTRTIVHYSNPLVFSGIEAIQKO-----SCESN----- 853
QY 841 SGALKSTSGSINHPIFPENSEVTFNITFPDVDSKASLGN-KLLLKANVTSENWPRNKT 899
DB 854 ----HNITKGVGPPFLRGEMVTFKILFOFNITSYLMENVTIYLSATSDSEEPETLSDNV 909
QY 900 FOLELPKAVYVMTSHGVTSTKYLNFTASNTSRVMQYOVSN-----LGORS----- 949
DB 910 VNISIPKVEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNFINFYLRKSGSP 968
QY 950 ---LPISLVF-----LVPVRLNQTVINDRPQVTFSENLSSTCHTK----- 987
DB 969 MPBLKLSIFPNNTSGNPVLYPTGLSS-----SENANCRPHIFEDPFSINSK 1017
QY 988 RLPSSHDFLAELRKAPVNGSIAVCORIQCDIPFFGQIE-----EFN 1029
DB 1018 KMTSTD----HLKRGITLDCNTCKFATITCNLTSSDIQVNSLILWKPTFKSYFSSLN 1074
QY 1030 ATLKGSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQGAFAVRSQETKVEPPEVN 1089
DB 1075 LTRGEL-----RENASLVUSSN-----QKRELAIOIKDGLPG 1110
QY 1090 PLPL--IVGSSVGLLLALLALITAAALYKLGFFKRYKDMNSE 1128

DB 1111 RVPLWVILLSAPAGLLMLLILALWKITGFFKRPULKQMEK 1151
RESULT 10
ITAH HUMAN
ID ITAH HUMAN STANDARD; PRT; 1189 AA.
AC O9UKX5; O9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITG11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
human integrin alpha11 subunit (ITG11).";
RL Genomics 60:179-187(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kuschel-Gullberg M., Sejersten T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
Integrin. A collagen-binding, i domain-containing, beta(1)-associated
integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
RP SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF109681; AAF01258.1; -
DR EMBL; AF137378; AAD51919.2; -
DR EMBL; AL359064; CAB94392.1; -
DR HSSP; P17301; 1A0X
DR Genew; HGNC:61336; ITG11.
DR MIM; 604789; -
DR GO; GO:0008305; C:integrin complex; TAS.

DR GO: GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO: GO:0005518; F:collagen binding activity; TAS.
 DR GO: GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO: GO:0007517; P:muscle development; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRIN.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.
 DR PROSITE; PS0234; VWF_A; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 1189
 FT DOMAIN 23 1142
 FT TRANSMEM 1143 1165
 FT DOMAIN 1166 1189
 FT REPEAT 38 94
 FT REPEAT 102 163
 FT DOMAIN 167 345
 FT REPEAT 359 420
 FT REPEAT 422 475
 FT REPEAT 477 537
 FT REPEAT 539 598
 FT REPEAT 601 653
 FT DOMAIN 1154 1162
 FT DOMAIN 1174 1177
 FT CA_BIND 488 496
 FT CA_BIND 551 559
 FT CA_BIND 613 621
 FT DISULFID 76 83
 FT DISULFID 121 139
 FT DISULFID 129 159
 FT DISULFID 659 668
 FT DISULFID 674 729
 FT DISULFID 781 787
 FT DISULFID 881 893
 FT CARBOHYD 82 82
 FT CARBOHYD 95 95
 FT CARBOHYD 291 291
 FT CARBOHYD 331 331
 FT CARBOHYD 358 358
 FT CARBOHYD 449 449
 FT CARBOHYD 462 462
 FT CARBOHYD 528 528
 FT CARBOHYD 642 642
 FT CARBOHYD 694 694
 FT CARBOHYD 857 857
 FT CARBOHYD 894 894
 FT CARBOHYD 973 973
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1040 1040
 FT VARIANT 433 433
 FT VARIANT 524 524
 FT VARIANT 972 972
 FT VARIANT 1003 1003
 FT VARIANT 1030 1030
 FT VARIANT 1094 1094
 SQ SEQUENCE 1189 AA; 133609 MW; 6030308AA4CD52 CRC64;

Query Match 18.4%; Score 1084.5; DB 1; Length 1189;
 Best Local Similarity 28.2%; Pred. No. 4.9e-66;
 Matches 349; Conservative 210; Mismatches 500; Indels 179; Gaps 48;

QY 1 FNLDTENAMTFOENARG-FGQSVVQ--LOGSR-VVGAPOEIVAAQNRGLSQCDVSTGS 56
 DB 23 FNDTRKPRVIFGSRTAFFGYVQQHDISGNKWLVGAPLETNGYKQTKGDKVYKCPVIHGN 82
 QY 57 CEPRL-----QVPVEAVNMSLGLSLAATSPFPOLLACGPTVHTQCSNTVYKGLCLFLP 110
 DB 83 CTKLNLRVTLGNVSRKDNMRGLSLATPNKDNSFLACSPLSWSECGSSYYTTGNCRSRV 142
 QY 111 GSNLRQOPQKPFPEARLRCQPEDSDIAFLIDGSGSIIPHDFRRMKWVST---VMEOLKK- 166
 DB 143 NSNFRSKTAP-ALQRC-QTYMDIVLDGNSIYP-----WVEVQHFILNLLKXF 192
 QY 167 ----SKTLFSLMQYSEEPRIHFTFKFQNNPNRSLIKBITQLLG-RTHATATGLRKVRE 221
 DB 193 YIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDVEAASHIEQRGGTETRTAFGIEPARSE 252
 QY 222 LFNITNGARKNAFKILFLTDGEKFGDPLGYEDVIPELDRGVIRVIGV-----GDAPR 276
 DB 253 AFO--KGRKGAKVMIVITDGESHDP--DLEKVIQOSRDNVTRYAVAVLGYNNRRGIN 309
 QY 277 SEKSROELNTVASKPRDRHVFOINNPEALKTIONQLREKIFAIEGTOTGSSSSFEHMSQ 336
 DB 310 PETFLNEIKYIASDDPKHFFNVTDAAUKDIVDALGDRIFSLGEGTNK-NETSGLEMQ 368
 QY 337 EGFSAAITSNGLPLSTVGSYDMAGVFLYTSKEK-----STFINMTRVDSMDNDAYLGYA 391
 DB 369 TGFSSHVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKNHGAFLGYT 428
 QY 392 AAILLRNV-QLVLGAPRYQHIGLVAMF-RONTGMWESNANVKGQIGAYFGASLCSVD 449
 DB 429 VTSWSSROGRVVVAGAPRNFHTGKVFLETHMNRSLTIHQAMRGOQIGSYFGSEITSD 488
 QY 450 VDSNGTDLVLGAPHYEQTR-GGOVSVCPPLRGORARWQCDVAVLGEQGPWGRFGAA 508
 DB 489 IDGGVTDVLLVGAPMYFNEGRERGKYVTEL---RQNRPVYNGTKDSHSYONARFGSS 545
 QY 509 LTVLGVDNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYF 568
 DB 546 IASVRDLNQDSYNDVVVGAPLEDNHAGAIYIPHGFRGS-ILKTPKQRIASELATGLQYF 604
 QY 569 GQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQPLVRKAIEMFNPREVARNVF--ECNDQ 626
 DB 605 GCSIHGQDLINEDGLIDLAVGALGNVILMSRPVQINASLHPFSKI--NIFHRDC--- 659
 QY 627 VVKGEAGEVRVCL-----HVQKSTRDLREGQIQSVVYTDLALDSGRPHSR 673
 DB 660 ----KRSGRDATCLAAFLCETPIFLAPHFOTITVG-----IRYNATMDERRYTTPR 705
 QY 674 AVFNETKNS-TRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNFSLVGTPLSAFGN 732
 DB 706 AHLDEGGDRFTNRVAVLLSSGOELCERINFHVLDTADYVYKPTVFSVEYSLEDP-----D 758
 QY 733 LRPVLADAORLTALFPFKKNGCNONI CODDL----- 765
 DB 759 HGPMLDDGWPTTLRVSVYFPWNGCNEDEHCVPLDVLDAARSDLPATAMEYCORVLRKPAQDCS 818
 QY 766 SITFSPMSLDCLVVGGPREFNVTVVRNDCGDSYRQVTFEFPDLDSYRKVSTLQNRSQ 825
 DB 819 AYTLSDFTTVFIIESTRQRVAVAEATLENRGENAYSTVLNISQANLQF--ASLIQKEDSD 876
 QY 826 RSNRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKAN 885
 DB 877 GS--IECVNEER-----RLQKQVCNVSYPFFRAKAKVAFRLDDEF-SKSIPLHLELELA 928
 QY 886 VTGNNNPRNTKTE--FQLELPVKYAVVMVVTSHGVSTKY---LNFTAS--ENTSRVMQH 938
 DB 929 AGSDSNRDTKEDNVNAPLRFHLKYEADVLFTRSSLSLHYEVKLNSSLERYDGTGPPFSC 988
 QY 939 QYQVSNLQO---RSLPISLVLPVRLNQTVIWRDPOVTFSENLSSTFC----HTKRLPS 991
 DB 989 IFRIQNLGLPFIHGMKMITIATRSNGNLLKLRLDFT-DEVANTSCNIGWNSTEYRPT 1047

Qy	992	HSDFLAELRKAPVNVCSIAVCRIQCDDIPFGIQEEFNATLKGNSLSPDWY-----IKTSH	1046
Db	1048	PVE--EDLRRAPQLNHSNSDVSNICNIRLVP-NQEIFHLGNL---WLRSLKALKYKS	1101
Qy	1047	NHLIIVSTAELFNDSVFTLLPGQGAIVRSQTEKTVPEFVPN-----PLPIIVGSSVG	1100
Db	1102	MKIMVNAALQRQH-SPF-----IFREEDPSRQIVFISKQEDMQVFIIVGSTLG	1152
Qy	1101	GLLLALITAAIYKLGFEEK--ROYKDMMSGGPPGAEP	1136
Db	1153	GLLLALVILVALWKLGFRRARRR-----PGLDP	1183

RESULT 11	
ID ITA2 BOVIN STANDARD; PRT; 1170 AA.	
AC P537IO;	
DT 01-OCT-1996 (Rel. 34, Created)	
DT 01-OCT-1996 (Rel. 34, Last sequence update)	
DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)	
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).	
DE ITGA2.	
GN Bos taurus [Bovine].	
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC Bovidae; Bovinae; Bos.	
OX NCBI_TaxID=9913;	
RN [1]	
RX SEQUENCE FROM N.A.	
RY MEDLINE=94193647; PubMed=7511592;	
RA Kanata T., Fuzon W., Takada I.;	
RT "Identification of putative ligand binding sites within I domain of	
RL integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";	
RL J. Biol. Chem. 269:9659-9663(1994).	
CC -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,	
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT	
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN	
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER	
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE	
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED	
CC EXTRACELLULAR MATRIX.	
CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2	
CC ASSOCIATES WITH BETA-1.	
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.	
CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS	
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.	
CC -I- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.	
CC -I- SIMILARITY: Contains 1 VWFA domain.	
CC -I- SIMILARITY: Contains 7 FG-GAP repeats.	
CC -----	
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to licensee@isb-sib.ch).	
CC -----	
DR EMBL; L25886; AAB59255.1; --	
DR PIR; I45914; I45914.	
DR HSSP; P17301; IA0X.	
DR InterPro; IPR000413; Integrin_alpha.	
DR InterPro; IPR002035; VWF_A.	
DR Pfam; PF01839; FG-GAP; 3.	
DR Pfam; PF00357; Integrin_A; 1.	
DR Pfam; PF00092; vwa; 1.	
DR PRINTS; PR01185; INTEGRINA.	
DR PRINTS; PR00453; VWFADOMAIN.	
DR SMART; SM00191; Int_alpha; 5.	
DR SMART; SM00327; VWA; 1.	
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR PROSITE; PSS0234; VWFA; 1.	

Db 370 VGFSAEYSPQNNILMLGAVGAYDWSGTVQKTPHGLIFSKQAFQILQDRNHSYLGYS 429
Qy 392 AAILNRVUSVLVGLAPRYQHIGLVAMFRONTGWESNANV-----KGTQICAYGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGOIYLSYN-----ENGNNVTVIQORGQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDLVLGAPHYEQTR--GGQVSCVPLPRQORARWQCDVLYGEOQOPWG 503
Db 485 CAVDVNKDTITDVLVGLAPRYMNDLKEGRVFLFTIKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAYVLFHGTSGSGISFSPHSQRIAGS--KL 561
Db 542 RFGSAIAALSINDMGDNFVIGSPLNQNAGVYVYNGHEGM-IRLRYSQKILGSDRAF 600
Qy 562 SPRLQYFGSLGQDITMGLVDLTVGAQHVLLLRSPQVLVAKMEINPREVARNVF 621
Db 601 SSSLQYFGRSLDGYDLNGSDITDVSVGAFGQVQVQLWSQSIADVDSATPTPKI--TLL 658
Qy 622 ECNDQVYKREAGVRVCLHVQKSTRDRLRREGQIQSVTVYDLALD----SCRPHSRAVEN 677
Db 659 NKNAEI-----KLKLCF-----SAKRPINQNNQVAIVYNTIDEDQFSSRVISRLFK 707
Qy 678 ETKNSTRQTOVLGLTQCB--TLKQLPNCIEDPVSPIVLRNFSL--VGTPLSAPGNL 733
Db 708 ENNERCLQTMIVSQAKRCSEYIIHQEPS---DIISPLNLCMNISLENPGT-----756
Qy 734 RPVLAEADQRLFTALFPKXKNCNDNICODLSITF----SPMSLDCVLVGGPREFNVTV 789
Db 757 NPALAEAYSETVKVFSIPFKDKCGDDGVCISDLVLNVQOLPATOQOPFIVSNQMKRLTFSV 816
Qy 790 TVRNDGDSYRTOVTFFPFLDLSYRKVSTLQNRQSQRSLRACASAST-EVSGALKSTS 848
Db 817 QLKVKESAYNTEIUVDFSENLF-----ASWSNPVDGTEVTQIASQSKSVT 864
Qy 849 CSINHPIPPENSEYTNFTDVSQKASLGNKLLKANVTSENMPRTNKTQFQLELPVKY 908
Db 865 CNVGPALKSKQQTFTINFDFNLQ-NLONQASISFRALSQSEENMADNSVNLKLSILY 923
Qy 909 AVYVMTSHGVSTKYLNFASENTSRVMOHQYQVSNLQOR-----SLPLSLFLV 958
Db 924 DAEIHT-RSTNINIFYSLDGNSSVV-HSFE--DIGPKFIFSIVKVTGSGVPVSMA---976
Qy 959 PVRLNQTVMIDRPOVTRSEN--LSSTCHTKRE-----RLPSSHDFLAE- 998
Db 977 -----SVIIHPQVTKDKNPLMYLTGVHTDQAGDISCEAINPLKIGTSSSVSFKSEN 1030
Qy 999 LRKAPVNVCSIAVCORIQCIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEI- 1057
Db 1031 FRHIKELNCRATSCSNINCMWLRDLQVKEGYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNSVFTL-----LPGQAFVRSQTKTEKPEF-VNPLPLIVGSSVGLLILALITA 1110
Db 1091 TYNPOIYVIEENTVTIP-----LTIMPKHEKVEVPTGVIVGSIAGIILLALVA 1140
Qy 1111 ALYKLGFFKPKYKDM 1125
Db 1141 ILWKLGFPRKRYEKM 1155

RESULT 12
ITA2 MOUSE
ID ITA2_MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Euthalia; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RNA SEQUENCE FROM N.A.
RP STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RC MEDLINE=94363406; PubMed=8081889;
RX Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologous supports collagen and laminin adhesion but
not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
ASSOCIATES WITH BETA-1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
EMBL; Z29987; CAA82877.1; --
DR EMBL; X75427; CAA53178.1; --
DR PIR; S41142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR004113; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178
FT DOMAIN 27 1129
FT TRANSMEM 1130 1151
FT DOMAIN 1152 1178
FT REPEAT 42 100
FT REPEAT ? ?
FT DOMAIN 185 375
FT REPEAT ? ?
FT REPEAT 431 483
FT REPEAT 485 546
FT REPEAT 548 607
FT REPEAT 612 664
FT CA_BIND 496 504
POTENTIAL.

FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GFPR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;
Query Match 18.0%; Score 1057; DB 1; Length 1178;
Best Local Similarity 27.8%; Pred. No. 3.7e-64;
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;
QY 1 FNLDENAMTFQ--ENARGFQSYVOL---QGRVVVVGAPQEIIVAAQNRGLSYQC--DYST 54
DB 27 YNVLPGAKIFSGPSSEQFGYSVQQLTNPQGNWLLVSGSPGPPENRMDGVYKCPVDLPT 86
QY 55 GSCEPIRLQ-----VPEAVNMSGLSLAATSPQLLACPTVHQTCSENTYVKGILC 107
DB 87 ATCEKLNQNSASISNVITKTNMSLGLTLTRNPGTGGLTCGLPLWAHQGNQYVATGIC 146
QY 108 FLFGSNLRQPOQ---KFPALRGCPQEDSDIAFLIDGSGIIPHPDRFMKENVSTMEOL 164
DB 147 ----SDVSDPDLTGSFAVQACPSL--VDVVVVCDESNIYP--WEAVKNFLKFPVTL 199
QY 165 K--KSKTLFSLMOYSBEFRIHFTKFEQNNPRSLIKPITQLLG--RTHATGLRKVRE 221
DB 200 DIGPKKTQVALIQYANEPRIIFNLNDFETKEDMVQATSETRQHGGLTNTFRAIEFARDY 259
QY 222 LFNITNGARKNAPKILFLTDGKFGDPLGYEDVIPELDRGVIRVIGV-----GDAPR 276
DB 260 AYSQTSGRGPGATKVMVVTVDGESH--DGSKLKTVIOCCNDDELTRFGIAGVLYGNRLALD 318
QY 277 SEKSRELANTVASKPRDRHVQFNFEALTKIQNLQREKIFATEGTQTCSSSFHEMSQ 336
DB 319 TKMLKEIKAIATPTERYFFNFVADAEALKEKAGTLGEQIFSEGTVQG--GDNFQMEMAQ 377
QY 337 EGFSA--AITSNGPLSLTVSGYDWAGGVFLYTSKEKSTFINMT--RVDSDMN--DAYLGYA 391
DB 378 VGFSADYAPQNDILMLGAGFDMWGTGLVQETSHKPVIPKQAFDQVLDQRNHSFLGYS 437
QY 392 AAILRNVRQSLVGLAPRYQHIGLVAMFRONTGMESNANV-----KGTQIGAYFGASLCS 447
DB 438 VAAISTEDGVHVFAGAPRANYTCQIVLYSNK---QGNVTVIQSHRGGDQIGSYFGSVLCS 494
QY 448 VVDVDSNGSTDLVLIGAPHYVEQTR--GGQVSVCLPRGQARWOCDAVLYGQGGQWGRF 505
DB 495 VVDVDDKTIITDVLVLGAPTYTNDLKKKEGVLYFTITKGLNQHQ----FLEGPEGTGNARF 551
QY 506 GAALTVLGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564
DB 552 GSAIALSDINMGDFNDVIGSVENENSGAVIYNGHOGT--IRTKYSQKILGNSGAFPR 610
QY 565 -LQYFGQSLGGQDLTMDGLDVLTVGAQGHVLLRSQPVLRKVAIMEFNPVARNVFEFC 623
DB 611 HLQFFGRLDGVGLDNGSDITDVSIGALGVQLWQSIADVAIEALFTP----- 660
QY 624 NDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVTYDLALD-----SGRPHSRVAFNET 679
DB 661 -DKITLLNKDAKITLKLCPRAEFPAGQNNQV--AILFNMTLDADGHSHSRVTSRGVFRN 717

QY 680 KNSTRQTOVLGLTQTCTET--LKQLPNCIEDPVPISVILNFSLVGTPLSAFGNLRPVL 737
DB 718 SERFLQKMMVNEVQKCEHHSIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
QY 738 AEDAQRLLFTALFPPEKNGNDNICODDLISL-----TFSEMSLCLVVGPPREFNVT 788
DB 769 EAYSETVKVFSIPFYKEGSGDGCISDLILDVQQLPAIQTSF-----IVSNQNKRLTFS 823
QY 789 VTVNRDGEDSYRTQVTFEPLDLSYRKVSTLQNRQSRQSWRLACASST--EVSGLKST 847
DB 824 VILKNRGESAYNTVVLAEFSENLF-----ASFMPVDGTETCEVGSQSKSV 871
QY 848 SCSINHPFPENSEVTFNITFDVSKASLGNKLLKANVTSENMPRTNKTE--FQLELP 905
DB 872 TCDVGYPALKSEQQVTFTFINFDNLQ--NLQAAINFOAFSESO--ETNKADNSVSLTIP 928
QY 906 VKYAVYVMVTSHGVSSTKYNLFTASENTSRVMQHVQVSNLGR-----SLPISLV 955
DB 929 LYDAELHLT--RSTNINFEISSDENAFSVIK---SVEDIGPKFISLKVTAGSAPVSWA 984
QY 956 FLV-----PVRLNQTVIMDRPOVTF--SENLS 980
DB 985 LVTHIHPQYTKENPLLYLTGQTQDAGDISCTAEINPLKLPHTA-----PSVSFKENFR 1040
QY 981 STCHTKERLPSSDPLAEELRKAPVNVVCSIAVCQRIQCDDIPFGIQEEFNATLKNLSFDM 1040
DB 1041 ---HTKE-----LDCRTTSCNITCWLKOLHMAEYFINVTRVMNRT 1080
QY 1041 YIKTSHNLLIVSTAEILFNDSVFTLLPQCGAFVRSQETKVEPEVENPLPLIVGSSVG 1100
DB 1081 FAASTFQVQVTAABEIDTHNPQLEBENAVTIPLMIMKPKTEKAEVPT--GVIIGSTIA 1138
QY 1101 GILLALITAAALYKLGFFKQYKDM 1125
DB 1139 GILLALLANTAGLWKLGFPRQYKDM 1163
RESULT 13
ID ITAL RAT STANDARD; PRT; 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=90338125; PubMed=2380249;
RX Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch F., Carbonetto S., Reichardt L.F.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen."
RL J. Cell Biol. 111:709-720(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliansky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alphabeta integrin I-domain: insights into
integrin I-domain function."
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

Db 877 CESN-----ONITCRVGYPLRAGETVTFKIIQFNTSHLSNAIHLIHSATSDSBE 927
QY 891 NMPTNTKTEQLEPLVKYAV-----YMWVTSHGVS-----KYLNTASENTSRVMOHQ 939
Db 928 PLESNDNEVNIPIKVEYGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFP 985
QY 940 YQVSNIGQSRSLP---ISLVP-----LVEVRLNQTVMD-----RP----- 971
Db 986 YTIKRGHFMPELQLSIFPNLTADGYPVLYPIG-----WSSSDNVNCRPSRSLDPFG 1039
QY 972 -----QVTPS-----ENLSSTCHTKERLPSPSHDSFLAELRKAPVNCISAVCQRIQ 1016
Db 1040 INSGKKWTISKSEVLKRGITQDCSSTC-----GVATITCSLLPSDLSQ 1082
QY 1017 CDT-----PFFGIGBEF---NATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLL 1067
Db 1083 VNVSLLLWKPTF-IRAFSSSNTLTGELK-----SENSSTLSSN----- 1123
QY 1068 POGAFVRQOTETKVPFEPVNPPLP--IVGSSVGLLLALITAAALYKLGFFKRYQKDM 1125
Db 1124 -----RKRELAQISKDGLGRVPLWILLSAFAGLLLLMLLILALWKIGFFKRLPKKK 1177
QY 1126 MSE 1128
Db 1178 MEX 1180

RESULT 14
IT2A2_HUMAN STANDARD; PRT; 1181 AA.
AC P17301;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
RC TISSUE=Endothelial cells;
RX MEDLINE=89308879; PubMed=2545729;
RA Takada Y., Hemler M.E.;
RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
RT (platelet GPIa): homology to other integrins and the presence of a
RT possible collagen-binding domain";
RL J. Cell Biol. 109:397-407(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
RX MEDLINE=98019223; PubMed=9353312;
RA Embley J., King S.L., Bergelson J.M., Liddington R.C.;
RT "Crystal structure of the I domain from integrin alpha2beta1";
RL J. Biol. Chem. 272:28512-28517(1997).
RN [4]
RP VARIANT HPA-5 (BR).
RX MEDLINE=94043762; PubMed=7901236;
RA Santoso S., Kalb R., Walka M., Kiefel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet alloantigens Br(a) and Br(b) are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2)";
RL J. Clin. Invest. 92:2427-2432(1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10744142;

RA Kroll H., Gardemann A., Fechter A., Haberboesch W., Santoso S.;
RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction";
RL Thromb. Haemost. 83:392-396(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NAIT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1; -;
DR EMBL; AF512556; AAM34795.1; -;
DR PIR; A33998; A33998
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MIM; 192974; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; F:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 1181 INTEGRIN ALPHA-2.
FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1133 1154 POTENTIAL.
FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1155 1161 INTERACTION WITH HPS5.
FT REPEAT 45 103 FG-GAP 1.
FT REPEAT 7 7 FG-GAP 2.
FT DOMAIN 188 378 VWFA.

```
FT REPEAT 378 433 FG-GAP 3.
FT REPEAT 434 486 FG-GAP 4.
FT REPEAT 487 549 FG-GAP 5.
FT REPEAT 551 610 FG-GAP 6.
FT REPEAT 615 667 FG-GAP 7.
FT CA_BIND 499 507 POTENTIAL.
FT CA_BIND 563 571 POTENTIAL.
FT CA_BIND 627 635 POTENTIAL.
FT SITE 1157 1161 GFEKR MOTIF.
FT DISULFID 83 92 BY SIMILARITY.
FT DISULFID 680 737 BY SIMILARITY.
FT DISULFID 789 795 BY SIMILARITY.
FT DISULFID 865 876 BY SIMILARITY.
FT DISULFID 1019 1050 BY SIMILARITY.
FT DISULFID 1055 1060 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 534 K -> E (IN ALLOANTIGEN HPA-5B;
dbSNP:1801106).
/FTid=VAR_003977.

FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT TURN 292 301
FT STRAND 302 303
FT STRAND 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 17.9%; Score 1054; DB 1; Length 1181;
Best Local Similarity 26.7%; Pred. No. 6e-64;
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLTENAMTQ-ENARGFGSVVQL----QGSVVVVGAPQEIIVANQSGSLYQC--DYST 54
Db 30 YNVLPEAKIFSGPSSEQFYAVQQFINPKGNWLLVGSPPWSGFFPENRMGDYKCPVDLST 89
Qy 55 GSCSPRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHTQCSNTYVKGLC 107
Db 90 ATCEKLNQSTSTSPNVTETMNTNMSLGLITRNMGTTGGFLTCGPLMWAQQCNQYVTTGVC 149
Qy 108 FLFGNLRQQPKPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMYQLK-- 165
```

Db 1142 GILLIALVALWKLQFFKKYEM 1166

RESULT 15

ITAG HUMAN STANDARD; PRT: 1167 AA.

AC 075578; Q9UH28;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-10 precursor.

GN ITGA10

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Articular chondrocytes;

RX MEDLINE=98352078; PubMed=9685391;

RA Camper L., Hellman U., Lundgren-Aakerlund E.;

RT "Isolation, cloning, and sequence analysis of the integrin subunit

RT alpha10, a beta1-associated collagen binding integrin expressed on

RT chondrocytes.";

RL J. Biol. Chem. 273:20383-20389 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Endothelial cells, and Heart;

RX MEDLINE=20169197; PubMed=10702680;

RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,

RA Wang S.-X., Langley R., Krissansen G.W.;

RT "The integrin alpha10 subunit: expression pattern, partial gene

RT structure, and chromosomal localization.";

RL Cytogenet. Cell Genet. 87:238-244 (1999).

CC - FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10

CC - ASSOCIATES WITH BETA-1.

CC - SUBCELLULAR LOCATION: Type I membrane protein.

CC - TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN

CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.

CC - DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC - SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC - SIMILARITY: Contains 1 VWFA domain.

CC - SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR ENBL; AF074015; AAC31952.1;

DR ENBL; AF112345; AAF21944.1;

DR ENBL; AF172723; AAF61638.1;

DR HSSP; P17301; IAOX

DR Genew; HGNC:6135; ITGA10.

DR MIM; 604042; -

DR GO; GO:0008305; C: integrin complex; TAS.

DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.

DR GO; GO:0005518; F: collagen binding activity; TAS.

DR GO; GO:0007160; P: cell-matrix adhesion; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR020335; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; vwa; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.

DR PROSITE; PS00234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; Repeat; Calcium; Magnesium.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 1167 INTEGRIN ALPHA-10.

FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1123 1145 POTENTIAL.

FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).

FT REPEAT 38 97 FG-GAP 1.

FT REPEAT 7 7 FG-GAP 2.

FT DOMAIN 167 350 VWFA.

FT REPEAT 365 427 FG-GAP 3.

FT REPEAT 428 482 FG-GAP 4.

FT REPEAT 483 545 FG-GAP 5.

FT REPEAT 546 605 FG-GAP 6.

FT REPEAT 608 660 FG-GAP 7.

FT DOMAIN 1134 1140 POLY-LEU.

FT CA_BIND 494 502 POTENTIAL.

FT CA_BIND 558 566 POTENTIAL.

FT CA_BIND 620 628 POTENTIAL.

FT DISULFID 76 86 BY SIMILARITY.

FT DISULFID 666 675 BY SIMILARITY.

FT DISULFID 681 736 BY SIMILARITY.

FT DISULFID 789 795 BY SIMILARITY.

FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 844 844 I -> L (IN REF. 2).

FT CONFLICT 909 909 G -> V (IN REF. 2).

FT CONFLICT 926 926 E -> D (IN REF. 2).

SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;

Query Match 17.9%; Score 1051.5; DB 1; Length 1167;

Best Local Similarity 28.7%; Pred. No. 8.7e-64;

Matches 351; Conservative 199; Mismatches 491; Indels 183; Gaps 45;

QY 1 FNLDTENAMTFQENARG-FGQSVVOLQSGR-----VVVGAPEQETVAANQRLSYOC----- 50

DB 23 FNLDDEHPRLFGPPEAEFGYVLQHVGGQRLMLVGAPEWGSGRRGDVTRCPVGGAH 82

QY 51 -----DYSTG-SCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCS 99

DB 83 NAPCAKAGHLGDYQLGNSSHP-----AVNMHLGMSLLETGDDGGFMACAPLWSRACGS 134

QY 100 NTYVVKGLCFGLGSNLRQQPKFPEALRGCPQSDSDIAFLDGSGLIIPHDFRMKRW--V 157

DB 135 SVFSGGICARVDASFPQSGSLAPTAQR-CPTY-MDVVIVLDGNSNIYP-----WSEV 184

QY 158 STVMEOL-----KSKSTLFLSMQVSEEFRIHFTKFEQNNPNRSLIKPITOLLGR-TH 210

DB 185 QTFLLRLVKGLFIDPEIQVGLVQVGSFVHWSLGDFTKKEVVAANKLSRREGRETK 244

QY 211 TATGLRKVVRELFNITNGARKNAFKILFLLTGKEF-GDPLGYEDYVIPELDREGVIRYVI 269

DB 245 TAAQIMVACTEGFSQSHGGRPEARLLVVVTGDSHDEEL--PAALKACEAGRVTRYGI 302

QY 270 GV-GDAFRSEKS-----RQELNTVASKPPRDHVFQINFEALKTIQNLREKIPAIGTQT 324

DB 303 AVLGHYLRQRDPSPSFLREIRTIASDPDRFFFNVTDEAALTDIVDALGDRIFGLEGSHA 362

QY 325 GSSSPSEHEMSQEGFAAITSNGPLLSVTGSDVWAGGVFLYTSKEKSTFINTRVDS--- 381

DB 363 ENESSFGLMSQIGFSTHRLKDGILPGMWGAYDMGQSVLWLEGHRLFPFRMALEDFPP 422

QY 382 --DMNDAYLGYA-AAIILNRVQSLVGLAPRYQHIGLVAMFR-QNTGMWESNANVKGTQI 437

Db 423 ALQNHAYLGYSSMLRGRRLLFLSGAPRFRHRGKVIQFQKKGAVRVAQSLQEQI 482
Qy 438 GAYEGASICSVDVDSNGSTDVLIGAPHYY--EOTRGQVSVCPPLPRQARWQCDVLY 495
Db 483 GSYFGSELCLPDLTDRDGTDLVLAAPMFQPNKGTGRVVYLV--GQOSLLTLQGTLO 540
Qy 496 GEOQPMWRFGAALTVLGDVNGDKLTDAIGAPGEEDNRGAVLFGHTSGSGISPSHSQR 555
Db 541 PEPQD-ARFGFANGALPDLNODGFADVAVGAPLEDGHOGALYLYHGTQ-SGYRPHPAQR 598
Qy 556 IAGSKLSPRLOYFQSGISGGODLTMDGLVDLTGAAQGHVLLLRQPVLRVKAIMEFNPRE 615
Db 599 IAAASMPHALSYFGRSVDGRDLDDGLVDVAVGAQGAAILLSSRPVHLTPSLEVTPQA 658
Qy 616 VARNVFCNDQVVKGEAG--EVRVCLHVQKSTRDLREGIOQSVVTYDIALDSGRPHSR 673
Db 659 ISVVQRCRR---RGQAVCLTAALCFQVTSRTFGRWDH---QFYMRFTASLDEWTAGAR 712
Qy 674 AVFNET--KNSSTRQTOVLGTLTOTCETLKLQPCIEDPVPSPVILRLNFSLVGTPLSAFG 731
Db 713 AAFDGSQORLSPRLRLSVG-NVTCQQLHFHVLDTSDYLRPVALTVTFALDNTTKPG-- 768
Qy 732 NLRVLAEDAQRFTALFPEKNCNGNDNICODDLSITFSFMSLDC-----LVGGPR 783
Db 769 ---PVLNEGSPTSIQKLVPSKDCGPDNECVTDLVQ---VNMDIRGSRKAPFVVRGRR 822
Qy 784 EFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEVSQA 843
Db 823 KVLVSTLTLENKENAYNTSLIIFSRNL---HLASLTPQR-ESPIKVECAPSA----- 872
Qy 844 LKSTSCSINHPIFPENSEVTENITFDVDSKASLG---NKLL-----LKANVTSENNMPT 895
Db 873 -HARLCSVGHVPFQTGAKVTLLEFEPSCSLLSQVFCGLTASSDSLERNGLQENTAQ 931
Qy 896 NKTEFQLELPVKYAVYVWVTSHGVSSTKYNFTASENTSRVMOHY----- 940
Db 932 -----SAYIQYEPH-----LLFSESTLHRYEHPYGTLPVGPGEFKTTL 972
Qy 941 QVSNLG---QRSPLSLVFLVP-----VRLNQTWIDRPQVTFSENLSSTCHTKER 988
Db 973 RVQNLGCYVSGLIISA--LLPAVAHGGNYFLSLSQVI-----TNNASCIVQNLTE 1021
Qy 989 LPSHSDFLAELRKAPVNVNCSTAVCQRIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNH 1048
Db 1022 PPGPPVHPEELOHTNRLNGSNTQCVVRCHLGQLAGTEVSVGLLRLVHNEFFRRAKFS 1081
Qy 1049 LLIVSTAEILFNDVFTLLPGGAFVRSQTEKVEPVPNPLPLIVGSSVGGLLLLALI 1108
Db 1082 LTVVSTFELGTEGVSVLQTEASRWSESLLEV-VQTRPILISLWILIGSVLGGLLLLALL 1140
Qy 1109 TAALYKLGFF-----KROYK 1123
Db 1141 VFCLWKLGGFFAHKKIPEEKREK 1164

Search completed: November 25, 2003, 14:17:30
Job time : 13.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARFGQ.....FRQYKDMSEGPGGAEPO 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriapi:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4371	74.3	1151	11	Q9J130
2	3891.5	66.1	1036	11	Q8CA73
3	3799.5	64.6	920	6	Q28984
4	3470	59.0	1169	4	Q81VA6
5	3310.5	56.3	1169	11	Q9QXH4
6	3236.5	55.0	1161	11	Q9QYF7
7	1523.5	25.9	1161	11	Q9WTV4
8	1513	25.7	1160	11	Q9R200
9	1395	23.7	1196	13	Q98TF1
10	1350.5	23.0	1086	4	Q96HB1
11	1344.5	22.9	1187	13	Q98TF0
12	1269	21.6	927	6	Q8HZV0
13	1154.5	19.6	1167	11	Q88340
14	1109	18.8	1167	11	Q88341
15	1041	17.7	1171	13	O42094
16	1032	17.5	1038	11	Q8BS01

17	1005	17.1	895	11	Q9WUF8	Q9wuf8 mus sp. itg
18	998.5	17.0	1160	6	Q8MKF4	Q8mkf4 felis silve
19	865	14.7	348	4	Q8TES5	Q8tes5 homo sapien
20	860	14.6	1332	5	Q9BPQ8	Q9bpq8 halocynthia
21	796	13.5	205	11	Q63001	Q63001 rattus norv
22	753.5	12.8	780	13	Q62271	Q62271 xenopus lae
23	738	12.5	823	4	Q8WY18	Q8wy18 homo sapien
24	686.5	11.7	823	11	Q8CE84	Q8ce84 mus musculus
25	669	11.4	1032	11	Q61989	Q61989 mus musculus
26	644.5	11.0	1033	6	Q9BGU3	Q9bgq3 bos taurus
27	643	10.9	1036	11	Q91YD5	Q91yd5 mus musculus
28	628.5	10.7	257	11	Q8C270	Q8c270 mus musculus
29	619.5	10.5	1041	5	Q9UB90	Q9ub90 lytechinus
30	614.5	10.4	1041	5	O76378	O76378 lytechinus
31	580	9.9	1034	13	Q98TT7	Q98tt7 gallus gall
32	579.5	9.8	1054	5	Q9U6S1	Q9u6s1 strongyloce
33	550	9.3	1033	13	O42598	O42598 xenopus lae
34	534	9.1	1016	13	Q91779	Q91779 xenopus lae
35	530	9.0	974	11	Q924W2	Q924w2 rattus norv
36	529	9.0	1073	11	Q8CC06	Q8cc06 mus musculus
37	526	8.9	1047	6	Q9MZD6	Q9mzd6 bos taurus
38	525.5	8.9	1007	6	Q9GK48	Q9gk48 bos taurus
39	512.5	8.7	1034	6	O9TUN4	O9tun4 oryctolagus
40	509.5	8.7	1036	6	Q9TUN6	Q9tun6 sus scrofa
41	506.5	8.6	1012	11	O70304	O70304 mus musculus
42	491.5	8.4	1049	5	Q8SY51	Q8sy51 drosophila
43	475	8.1	833	5	Q9BPQ7	Q9bpq7 halocynthia
44	469.5	8.0	1036	6	Q9TU44	Q9tu44 canis famil
45	468.5	8.0	604	11	Q8BQ25	Q8bq25 mus musculus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
 ID Q9J130
 AC Q9J130
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr.; Zerrila K. Jr.;
 RT "Cloning of the rat cd11b cDNA sequence."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268593; AAF81280.1; -
 DR HSSP; P11215; 1BHQ
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWEA; 1.
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.3%; Score 4371; DB 11; Length 1151;
 Best Local Similarity 72.6%; Pred. No. 9.4e-317;
 Matches 826; Conservative 150; Mismatches 159; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARFGQSVVVLQGSRVVVCAPQEIIVAAHQSGSLYOCIDYSTGCEPI 60
 DB 17 FNLDTENPMTFQENASFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNRCDPI 76

```
Qy 61 LQVPPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCFEGSNLRQPOK 120
Db 77 LQVPPVAVNMSLGLSLAATTVPPOLLACGPTVHONCKENTYVNGCLYFGSNLRLRPQQ 136
Qy 121 FPEALRGCPQSDSIAFLIDSGSGSIIPHDFFRMKEWSTVMEQKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQESNIAFLIDSGSGSINTIDFQKMEFVSTVMDQFQKSKTFLSLMQYSEF 196
Qy 181 RIHTTFKEFQNNPRSLIKPITQLGRTHATGLRKVRVRELFTNGARKNAFKILFL 240
Db 197 RTHFTFNDKRNPPKSHVRPIRLQNGRTKTASGIRKVRVRELFTNGARKNAFKILFL 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVGDFAFRSEKSPQELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVIGVGDFAFRSEKSPQELNTVASKPPRDHVFQIN 316
Qy 301 NFPEALNTIQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPPLSTVGSYDWAG 360
Db 317 NFPEALNTIRNLQLEKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFTINMTVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFINTTRIDSDMDNDAYLGASAVISRRNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTREGGQVSCPL 480
Db 437 QNFGAWEPHTTIKSGQISYFGASLCSVDMADGNTNLIILGAPHYEKTREGGQVSCPL 496
Qy 481 PRGQARWQCDAVLGEGCQPGWRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-QARWQCEALHGDGQHPNGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 600
Db 556 HGASVASISTPSHSQRIAGARFSPGLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 615
Qy 601 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGQIOSVVT 660
Db 616 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGQIOSVVT 675
Qy 661 YDLALDGRPHSRVFNSTKSTRQTQVLGTTOTCETLKLQPLNCIBDPSPVILRLNF 720
Db 676 YDLALDGRPHSRVFNSTKSTRQTQVLGTTOTCETLKLQPLNCIBDPSPVILRLNF 735
Qy 721 SLVGTPPLSAFGLNLRVLAEDAQRLFTALPFEKNCNGNDNI CODDLSITFSPMSDCLVVG 780
Db 736 TLVGSPLASSRDLRPVLAEAQRIEFTAMPPEKNCNGNDNI CODDLSITFSPMSDCLVVG 795
Qy 781 GPRFNVTVVRNDCGDSYRQVTFPPFLDLSYRKVSTLQNRQSRWSLACESASTEV 840
Db 796 DSRDFVSVTLNRDGEDSYGKVTCTYPSGLSYRKVSASQNFQSKPWRVIAE-PSSE 854
Qy 841 SGALKASTCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 QGVLASTIWDINHPIFPANSEVTFNVTFDVDSASLGNKLLKVVVTSENNVSRDKTES 914
Qy 901 QLELPVKYAVNMTVSHGVSTKYLNFNTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 960
Db 915 QLELPVKYAVNMTVSHGVSTKYLNFNTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 974
Qy 961 RLNQTVIWDNRQVTFSENLSCTCHTKERLPSSHDFLAELRKAPVNCISIANVCORLOCDIP 1020
Db 975 QINKVTIWDNRQVTFSENLSCTCHTKERLPSSHDFLAELRKAPVNCISIANVCORLOCDIP 1034
Qy 1021 PFGIOEFNATLKGSLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1080
Db 1035 SPNSKEIFNVTLQGNLLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1094
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLALITALYKLGFFKQYKQKMMSEGGPPGABPQ 1137
Db 1095 KVEPEVNPPLIIVGSSVGGLLLALITALYKLGFFKQYKQKMMSEGGPPGABPQ 1151
```

```
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.",
RL Nature 420:563-573(2002).
DR EMBL; AK039444; BAC30350.1; -.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;
```

```
Query Match 66.1%; Score 3891.5; DB 11; Length 1036;
Best Local Similarity 66.0%; Pred. No. 4.8e-281;
Matches 751; Conservative 125; Mismatches 143; Indels 119; Gaps 2;
```

```
Qy 1 FNLDTENAMTFQENARFGQSVQLOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENARFGQSVQLOGSIVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76
Qy 61 LQVPPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCFEGSNLRQPOK 120
Db 77 LQVPPVAVNMSLGLSLAATTVPPOLLACGPTVHONCKENTYVNGCLYFGSNLRLRPQQ 136
Qy 121 FPEALRGCPQSDSIAFLIDSGSGSIIPHDFFRMKEWSTVMEQKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQESNIAFLIDSGSGSINTIDFQKMEFVSTVMDQFQKSKTFLSLMQYSEF 196
Qy 181 RIHTTFKEFQNNPRSLIKPITQLGRTHATGLRKVRVRELFTNGARKNAFKILFL 240
Db 197 RIHTFNDKRNPPKSHVRPIRLQNGRTKTASGIRKVRVRELFTNGARKNAFKILFL 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVGDFAFRSEKSPQELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVIGVGDFAFRSEKSPQELNTVASKPPRDHVFQIN 316
Qy 301 NFPEALNTIQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPPLSTVGSYDWAG 360
Db 317 NFPEALNTIRNLQLEKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFTINMTVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFINTTRIDSDMDNDAYLGASAVISRRNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTREGGQVSCPL 480
Db 437 ENFTWEPHTSIK----- 450
Qy 481 PRGQARWQCDAVLGEGCQPGWRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 600
Db 451 -----SQRIIGAHFSPGLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 498
Qy 601 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGQIOSVVT 660
Db 499 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGQIOSVVT 558
```

```
QY 661 YDIALDSGRPHSRVAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 559 YDIALDPGRSIRARFAFDETKNTRRRRTQVGLMKQKCTLLKLPDCVDDSVSPILRLAY 618
QY 721 SLVGTPLSFGNLRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVAG 780
DB 619 TLVGEPLRSFGNLRPVLAEDAORFFAMFPPEKNCNDNICODDLSITFSFMSLDCLVAG 678
QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWL-ACESASSTE 839
DB 679 GPQDFNMSVTLRNDGSDSYRTQVTVYPSGLSYRKDSASQNPITKKPMFKVPAESSSSE 738
QY 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTE 899
DB 739 GHGALKSTTWNHPIPPANSEVTFNITFDVDSHAGSFGNKLKLLKAIVASENNMSRTHKTK 798
QY 900 FQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHGYOVSNLGQSRSLPLSLVLP 959
DB 799 FQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHGYOVSNLGQSRSLPLSLVLP 959
QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 1019
DB 859 VQINNVTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 918
QY 1020 PFFGIOEENATLKNLSFPMYIKTSHNHLIYSTABILFNDVSFTLLPQOGAFVRSQTE 1079
DB 919 PSFNTQEIFNVTLLKGNLSFPMYIKTSHNHLIYSTABILFNDVSFTLLPQOGAFVRSQTE 978
QY 1080 TKVEPEVHPNPLVLIVSSVGGLLLALITAAIYKLGFFKRYQKMDMSEGGPPGABPQ 1137
DB 979 TKVEPEVHPNPLVLIVSSVGGLLLALITAAIYKLGFFKRYQKMDMSEGGPPGABPQ 1036

RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT NON_TER 920
FT SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.6%; Score 3799.5; DB 6; Length 920;
Best Local Similarity 78.8%; Pred. No. 2.9e-274;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;

QY 118 POKFPEALRGCPQEDSDIAFLIDSGSIIIPHFRMKEWSTVMEQLKSKTFLSLMOYS 177
DB 1 POKFPEALRGCPQEDSDIAFLIDSGSIIIPHFRMKEWSTVMEQLKSKTFLSLMOYS 60
```

```
QY 178 EBFRIHTFKEQNPNPRLSKPIKTOLLGRTHTATGLRKVYRELFTNTNGARKNAFKIL 237
DB 61 EDFYTHFTFNDPKRNPSPKLLVRP:ROLGRTHTATGIRKVVYRELFSKSGARENAFKIL 120
QY 238 FLTLTGKFGDPLGVYEDVTPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRPDHVP 297
DB 121 VVITDGEKFGDPLGVYEDVTPEDARKGVIRYVIGVDADFNSWKSRELNTIASKPCGDHVP 180
QY 298 QINNFEALKTIQNLQREKIFAIEGTQTSSTSSSFEHMSQEGFSAAITSNGLPSTVGSYD 357
DB 181 QVNNFEAVKTIQNLQREKIFAIEGTQTSSTSSSFEHMSQEGFSAAITSNGLPSTVGSYD 240
QY 358 WAGGVFLYTSKESKSTFINNTRVDSMDNDAYLYAAAIIILNRVQSLVLGAPRYQHIGLVA 417
DB 241 WAGGAFHLMPKORVIFINTTRVDSMDNDAYLYAVEVILRNOAQSLVLGAPRYQHTGLV 300
QY 418 MFRONTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQGVSV 477
DB 301 MFKQNSGAWEKWADIKGSIQISYFGASLCSVDVNRDGGSDLVIGAPHYETRGQGVSV 360
QY 478 CPLPRQARMQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLPQG-RAKMQCRVILCGEQGHPSRFGAALTALGDVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLL 597
DB 420 YLFHGTSELGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLL 479
QY 598 RSQPVLRKAIEMFNPREVARNVFCNDQVVKGEAGEVYVCLHVOKSTRDLRREGQIOS 657
DB 480 RSQPVLRKAIEMFNPREVARNVFCNDQVVKGEAGEVYVCLHVOKSTRDLRREGQIOS 539
QY 658 VVTYDIALDSGRPHSRVAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLR 717
DB 540 IITYDLALDPGRPHPRAVEETKNTRROTQVGLSRKCEHLALWLPDCVEDSVTVILR 599
QY 718 LNFSLVGTPLSFGNLRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLDC 777
DB 600 LNFSLVGTPLSFGNLRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLDC 659
QY 778 VVGCPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESAS 837
DB 660 VVGCPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESAS 719
QY 838 TEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
DB 720 TEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNK 779
QY 898 TEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHGYOVSNLGQSRSLPLSLV 957
DB 780 TEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHGYOVSNLGQSRSLPLSLV 839
QY 958 VPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCOR 1017
DB 840 VPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCOR 899
QY 1018 DIPFPGIOEENATLKNLSF 1038
DB 900 DIPFPGIOEENATLKNLSF 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```



```

RN      RC      RA      RL      DR      SQ
SEQUENCE FROM N.A.
TISSUE=Blood;
Strasbourg R.;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
EMBL: BC038237; AAH38237.1;
SEQUENCE 1169 AA; 128521 MW; A17B484FEFC79EB6 CRC64;

Query Match
Best Local Similarity 59.0%; Score 3470; DB 4; Length 1169;
Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

Qy      Db
1  FNLDENAMTFOENARFGOSVVOLOQSRVVVGAPOEIVAAANORGLSYQCDYSTGSCPEI 60
20  FNLOTEELTAFRVDSAGFGSVVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
61  RLQVPVEAVNMNLSGLSLAATTPPOLLACGPTVHOTCSENTYVKGLCFGLFSGNLRQOPQK 120
80  GLQVPPPEAVNMNLSGLSLAATTPSOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
121  FPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMKEWSTYMEQLKSKTFLSLMOYSEEF 180
138  LPVSRQECPRQEQDIVLIDGSGSISRNFATMNFVRAVISQORPSTQFSLMQFSNKF 197
181  RIHTEFEFONNPRSLIKPITOLLGRTHATGLRKVVRELFTNGARKNAFKILFL 240
198  QTHTEFEFRSSNPLSLASVHQLQFTYTATAIQNVVHRLFHASGARDAKILIVI 257
241  TDGKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDHFVQIN 300
258  TDGKKEGDSLDYKDVPWADAAGIIRVAIGVLAFQNRNWKELNDIASKPSQEHFKVE 317
301  NFEALKITQNLQRKIFAIETQGTGSSSPFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
318  DFDALKDQNLQKKEKIFAIETGTTSSSSFELEMAQEGFSAVFTPDGPVLGAVGFSWMSG 377
361  GVLVYTSKEKSTFTNMTVRDSDMDAYLGYAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
378  GAFLYPPNMSTFTINMSQENVDMRDSYLCGYSTELALWKGVSLVGLGAPRYOHTGKAVFT 437
421  QNTGMWESNANVKGTYGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQGVSVCP 480
438  QVSRQWRMKAETGTQIGSYFGASLCSVDVDSGSDTLVLIGAPHYYEQTGQGVSVCP 497
481  PRGQARQOCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
498  PRGWR-RWMCDAVLYGEOGHPGRFGAALTVLGVDVNGDKLTDVWIGAPGEKENRGAVYLF 556
541  HCTSGSGLSPSPSQRISAGSKLSPRLQYFGOSLGGGDLTMDGLVDLTVGAOQHVLRLSQ 600
557  HGVLGPSLSPSHQRIAGSLSRRLQYFGQALSQGGDLTQDGLVDLAVGARGQVLLLR 616
601  PVLRYKAIMFENPREVARNFECNDQVYKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
617  PVLWGVSMQFIPABIPSAFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDIQSSVT 676
661  YDLALDSRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
677  LDALDLPGRSLSPRATFQETKNSRSLSRVRLGLKAHCENFNLLPLSCVEDSVTPITLRLNF 736
721  SLVGNTPLSAAGNLRPLVAEDAORLFTALFPPEKNCNDNI CODDLSITFSFMSLCLVVG 780
737  TLVGLPLAFNLRPLVAADAQRYTASLPPEKNCADHICQDNLIGISFSPGLKSLVVG 796
781  GPREFNVTVVRNDEGDSYRTQVTEFFFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
797  SNLELNAEVMWNWNGEDSYGTTVTFSHPAGLSYRVVAEQKQGLRSLHLTCDSPVGV-- 854
841  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRNTKTEF 900
855  SQGTWSTSCRINHILFRGAQITFLATFDVSPKAVLGDRLLLTANVSENNTPTSKTTF 914
901  QLELPVKYAVYVMTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLQGRSLPISLVLVP 959
```

```

Db      Qy
915  QLELPVKYAVYVMTSHGVSTKYLNFTASEEESHVAMERYQVNNLQORDLPVSNFWVP 974
960  VRLNQTVIWDRPQVTFSENLSTCTKRLPSHSDFLAELRKAPVNVNCIAVCORIQCDI 1019
975  VELNQEAVMVDVEVSHQPNPSRCSSEKTAIPASDFLAHIQKNPVLDSCSIAGCLRFRCDV 1034
1020  PFGIOEEFNATLKGNSLDFWIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSOTE 1079
1035  PSFSVQEEELDTLKGNSLFGWVRQILQKVVSVVAEITFDTSVYSQLPQGEAFMAQTT 1094
1080  TKVPEPEVNPPLIVSSVGGLLLALITALYKLGFPKQYKQNMSE 1128
1095  TVLEKYKVHNPPLIVSSSIGLLLALITAVLYKVGFQYKQYKEMMEE 1143

RESULT 5
Q9QXH4 ID Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattiss F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAF23492.1; -.
DR HSP; P11215; IBIHQ.
DR MGD; MGI:96609; Itgax.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin.
KW SEQUENCE
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.3%; Score 3310.5; DB 11; Length 1169;
Best Local Similarity 56.6%; Pred. No. 1.3e-237;
Matches 645; Conservative 173; Mismatches 307; Indels 19; Gaps 7;

Qy      Db
1  FNLDENAMTFOENARFGOSVVOLOQSRVVVGAPOEIVAAANORGLSYQCDYSTGSCPEI 60
20  FNLOTEELTAFRVDSAGFGSVVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
61  RLQVPVEAVNMNLSGLSLAATTPPOLLACGPTVHOTCSENTYVKGLCFGLFSGNLRQOPQK 120
80  GLQVPPPEAVNMNLSGLSLAATTPSOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 138
121  FPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMKEWSTYMEQLKSKTFLSLMOYSEEF 180
139  PPTAQECPCQDQDIVLIDGSGSISRNFATMNFVRAVISQORPSTQFSLMQFSNKF 198
181  RIHTEFEFONNPRSLIKPITOLLGRTHATGLRKVVRELFTNGARKNAFKILFL 240
199  RVHTTFNFIKSTSPSLSLGSRVQLRGYTYTASAKHVITELFTTQSGARQDQATKVLVI 258
241  TDGKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDHFVQIN 300
```

```
Db 259 TDGRKQDNLSDVSPMAEASIIIRYAGVGKAFYNEHSKQELKAIASMPSEHYVFSV 318
QY 301 NFEALKTIONQOLREKIFALEGTQSSSFEHMSQEGFSAATISNGPLLSWTGSDWAG 360
Db 319 NFDALXDIENQLEKIFALEGTETPSSTFEHMSQEGFSAVFTDGPVLGAVGFSWSG 378
QY 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGVAIAIILNRNVQSLVGLGAPYQHIGLVAMFR 420
Db 379 GAFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIF 438
QY 421 QNTGMWESNANVKGTOIGYFGASCLSDVDNSGTDVLIGAPHYYEOTRGQVSVCLP 480
Db 439 QSRHWRPKSEVRGTQIGSYFGASCLSDVDNRDGDSTDLVLI GPHYYEOTRGQVSVCPM 498
QY 481 PRQARMOCDAVLYGEOQOPWGRFGAALTLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 499 P-GVGRWCGTTLHGEQHPWGRFGAALTLDVNGDLSADVAIGAPGEENRGAVYIF 557
QY 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPSORISASQIPRIQYFGQSLSGGQDLTRDGLVDLAVGSKRVLRLTR 617
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 618 PILRVSPVTHFTPAEISRSVFECQEAPEQTLSDATVCLHIHESPKTL--GDLRSTVT 675
QY 661 YDLALDSGRPHSRVNETKNSRRROTQVLGLTQTCETLKLQLPNCIEBPVSVILRLNF 720
Db 676 FDALDHGLSLSTRAIFKETKTRALTRVKTGLGNKHCESVKLLLPACVEDSVPTITLRLNF 735
QY 721 SLVGTPLSAGNLRPVLAEDAQRFTALFPFEKNCNDNICDDLSITFSFMSLCLVVG 780
Db 736 SLVGVPISSLQNLQPLAVDDQYFTASLPFEKNCGADHICDDLSVWFGFPLDKTLVVG 795
QY 781 GRPREFNVTVNRDGEDSVRTQVTFPPLDLSYRKVSTLQ-----NQRORSWR 829
Db 796 SDELELVDTVNSDGEDSVGTITLFPVPLGSLFRVAEQVFLRKEDQOQRGQSHLH 855
QY 830 LACESASSTEVSGALKSTCSINHPFENSEVTNITFDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSGLMSTSCSRHVIFRGSGQMTFLVTFDVPKABGLDRLLRLARVGE 913
QY 890 NNPRTNKTETFOLELPVKYAVYVTVSHGVSTKYLNTASE-NTSRVMQHOYQVSNLGR 948
Db 914 NNVPGTPTTFOLELPVKYAVYTMISSHDOFTKYLNFSTSEKETSVEHREQVNNLGR 973
QY 949 SUPISLVPLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVYVNC 1008
Db 974 DVPVSINFVPIELKGEAVW-TMVSHQPQNLPTQCYRNLKPTQFDLLTHMQSPVLDCS 1032
QY 1009 IAVCQRIQCDIPFFGQIEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDSVFTLLP 1068
Db 1033 IADCLHLRCDIPSLGLDDELFTLKNLSFGFISQTLQKVLULLSABEITFNTSVYSQLP 1092
QY 1069 GQAFVRSOTETKVEPFPENPLPLTVGSSVGLLLALLITALYKLGFFKQYKDMME 1128
Db 1093 QGEAFRAQTKTVLEMYKVNPPVPLVGVSSVGLLLALLITALYKAGFFKQYKEMLEE 1152

RESULT 6
Q0QVE7
ID Q0QVE7 PRELIMINARY; PRT; 1161 AA.
AC Q0QVE7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.0%; Score 3236.5; DB 11; Length 1161;
Best Local Similarity 57.5%; Pred. No. 4.4e-232;
Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLDTENAMTFOENARGFGQSVVQLGSRVWVGAPOEIVAAANQSGSLYOCDSYTGSCPEIR 61
Db 21 NLDVEEPIVFREDAAASFGQTVVQFGSRLVVGAPLEAVAVNOTGLYCAPATGWCQPIV 80
QY 62 LQVPVEAVNMSLGLSLAAATSPQLLACGPTVHTQCSENTYVYKGLCFLFGSNLRQOPQKF 121
Db 81 LRSPLVEAVNMSLGLSLVATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139
QY 122 PEALRGCEQESDIAFLIDGSGSIIPHPFRKKEWVSTVMEQLKSKTLFSLMOYSEBFR 181
Db 140 PASMPECPQEQEMDIAFLIDGSGSINQRDFAQKDFVKALMGEBFASSTLFLSMOYSNILK 199
QY 182 THFTPEKQNNPNRSLIKPITQLGRTHATGLRKVKVRELFNITNGARKNAFKTLFLLT 241
Db 200 THFTTEFNILDPQSLVDPIVQLQCLTYTATGIRTVMEELFHSKNGSRKSAKLLLVIT 259
QY 242 DGEKFGDPLGYEDVPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVFOINN 301
Db 260 DQKYRDPLEYSVDVIPAADKAGIIRYAGVDAFQEPALTELKELNTIGSAPPQDHVKVGN 319
QY 302 FEALKTIONQOLREKIFALEGTQSSSFEHMSQEGFSAATISNGPLLSWTGSDWAG 361
Db 320 FAALRSIQRLQEKIFAIEGTQSSSFEHMSQEGFSAATISNGPLLSWTGSDWAG 379
QY 362 VFLYTSKESKSTFINNTRVDSMDNDAYLGVAIAIILNRNVQSLVGLGAPYQHIGLVAMFRQ 421
Db 380 AFLYPNTRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIF 439
QY 422 NTGMWESNANVKGTOIGYFGASCLSDVDNSGTDVLIGAPHYYEOTRGQVSVCLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASCLSDVDNRDGDSTDLVLI GAPHYYEOTRGQVSVFP 499
QY 482 RGQARMOCDAVLYGEOQOPWGRFGAALTLDVNGDKLTDVAIGAPGEEDNRGAVYLF 541
Db 500 -GVGRWQCEATLHGEQHPWGRFGAALTLDVNGDNLADVAIGAPGEESRGAVYIF 558
QY 542 GTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSOP 601
Db 559 GASRLIEIMPSPQSVTGSQSLURLQYFGQSLSGGQDLTQDGLVDLAVGAGHVLRLSRLP 618
QY 602 VLRVKAIMEFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDLREGQIQSVVT 661
Db 619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGETVCLTVHKGSFDLL--GNVGSVRY 676
QY 662 DLALDSGRPHSRVNETKNSRRROTQVLGLTQTCETLKLQLPNCIEBPVSVILRLNFS 721
Db 677 DLALDPGRILSRAIFDETKNCTLTKRTKTLGLGDHCHETVKLLPDCVEDAVSPILRLNFS 736
```


Db 118 SYKTEPDSYVRKDDALLKHVHMLLTNTFGAINVYVATEVFRRELGARDPATKVLII 177
QY 239 LLTDGKFGDPLGVEDVIPLEDEGVIRYVIGVDARFRSEKROELNTVASKPRDRHFQ 298
Db 178 IITDGE--ATDSNIDAADK-----IIRYIIGIKHFKQTESQETLHKFASKPASFFVKI 230
QY 299 INNFEALKTIONLREKIFAIBGTQGTSSSFHEMSQEGFAAITSNGPLLSSTVGSYDW 358
Db 231 LDTFEKLKDLTFELQKIVYIEGTSKQDLTSFNMELSSSGISADLSRGAHVAVGAKDW 290
QY 359 AGGVF-LYTSKESKSTFINMTVDSDMNDAYLGAAA-IILNRVQSIVLGAPRYOHIGLV 416
Db 291 AGFLDLKADLQDDTFIGNELPTEVRAGLVGVTVTLPSRQKTSLLASGAPRYQMGVRV 350
QY 417 AMFR--QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQ 474
Db 351 LLFQEPGGGHWSQVQTIHTQIGSYFGELCGVDVDQGETELLIGAPLFYGEORGR 410
QY 475 VSCVPIPRQARWQCDV--LYGEOGQPMGRGAALTVLGDVNGDKLTDVAIGARGEED 532
Db 411 VFYI-----QRQLGFEVSELOQDPCYPLGRFCEAITALTIDINGDGLVDVAVGAPLEE- 464
QY 533 NRCAYVLFHGTSSGSLSPSHSRIAGSKLSPRLOYEQSLSGQDLTMGLVDLTVGAQG 592
Db 465 -QCAVIFNGRHG-GLSPQSPQRIEGTVLSGIGQWFGRSIHGVKDLGGLADVAVGAES 522
QY 593 HVLLLRASQVLRVKAIMEFNPREVARNVFCNDVV-KGKEAGEVRCVHLHVQKSTDRDLR 651
Db 523 QMIVLSRPVDMVTLMSPASIPVHEVECSYSTSKMKEGYNIICQI-KSLIPQF- 580
QY 652 EGOIQSVVYVDLALDSGRPHSAVFNETKNSTROTQVGLGTOTCBTLKLQLPNCIEDPV 711
Db 581 QGRLVANLTYTLQDGHRRRRGLFPGRRHRLRNIAVT-TSMSCDTSFHPFVPCVDLI 639
QY 712 SPIVLRLNLSL---VCTPLSAFGN-----LRPLAEDAQRLLFTALPPEKNCNDNICQ 762
Db 640 SPINVSLSLWEEEGTPDRQAGKQIPILRLPSLHSETWEI-----PFEKNGCDKKCE 694
QY 763 DLSLITFSMSLDCVLVGGPREFNVTVVRNDEGDSYRTQVTFEFLDLISYKRVSLQW 822
Db 695 ANLRVSFSPARSALALTAFAISLVELSLNLEEDAYWQLDLHFPFGLSFRKVELM--- 751
QY 823 RSORSRLACES--ASSTEVSGALKSTSCSINHPIPPENSEVTNTTFDVSASLGNKL 880
Db 752 KPHSQIPVSCPEELPEESRLISRAL--SCNVSSPIFKAGHSVALQWMTLVNSSWGDSV 808
QY 881 LKANYTSEN---MPTNKTEFQLELPVKYAVVWVTSHGVSSTKYLNFTASENTRVM 936
Db 809 ELHANYTCNNEDSULLEDNSATTI---IPILYPINILIQOQEDSTLYVSFTPKPKIHQV 865
QY 937 QHOYQV--SNLQORSIP-LSLVLPVRLNQTVIWRDPQVTFSENLSSTCHTK--ERLP 990
Db 866 KHYQVRIQPSIHDNIPTLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYEDLERLP 923
QY 991 SHSD--FLAELRKAIPVNVCSIAVCORIQCDIPFGIQEBFNATLKNLSFDWYIKTSHNH 1048
Db 924 DAAEPCLPGALPCPVV-----FRQELVQVIGTLELVGEIEAS-SM 964
QY 1049 LLIVSTAELFNDVSFTLLPGCAFVRSQTETKVEPFPVNPPLIVGSSVGGLLLLALI 1108
Db 965 FSLCSSLSISFNSKHFHLYGSNASL-AQVVMKVVDVYKQMLYLVLSVIGLGLLLLLI 1023
QY 1109 TAALYKLGFPKQYKDMWSEG-GPPGAEP 1136
Db 1024 FIVLYKVGFKRLNKEKMEAGRVPNGIP 1052

RESULT 11

Q98TF0

ID Q98TF0 PRELIMINARY; PRT; 1187 AA.

AC Q98TF0;

DT 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1; .
DR HSSP; P20701; ILPA.
DR InterPro; IPR001969; Asprotease site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 22.9%; Score 1344.5; DB 13; Length 1187;
Best Local Similarity 31.3%; Pred. No. 7e-91;
Matches 373; Conservative 209; Mismatches 453; Indels 157; Gaps 43;

QY 1 FNLDTENAMTFQENARG-FCQSVVQLQ-GSR--VVVGAPQEIIVANQORSLVQCDYSTGS 56
Db 32 FNIDTEHPLRFNGTPEDFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSC7ADLOS 88

QY 57 CSPIRLQVP---VEAVNMSLGLSLAATSPQLLACGPTVHTQTSNTYVVKGLCLFLGS 112
Db 89 CK--RLQPGSESVRFVFGNSAAVSSNAITS-----CSPYFAHECDGNSYLVNGVCYQFNS 140

QY 113 NLRQOPQKPEALRQCPQEDSDIAFLIDGSGSIIIPHFRMKEWSTVMEQLKSKTFLS 172
Db 141 SL-QAVSNFTAAAYQECSEKREVLVFLFDGSSMKAVEFDWKNFKIDVMMKLSNSSIKFA 199

QY 173 LMQYSEEFRIHFTKFEFQNNPNRSLIKPIITOLLGRTHATGLRKVVRELFN-ITNGARK 231
Db 200 AVQFSTEIRTVDFPDYQNGSABEKLME-RHMKSLTNTYKAINYYLVKNLVNSVSSGADP 258

QY 232 NAFKILFLLLTGCEKFGDPLGYED--VIPELDREGVIRYVIGVDAPFRSEKSRQELNVTAS 289
Db 259 NAKALVITD---GDPSDDNDYINILCDQNILRYIIGVG-----KVDLTTLTQLAA 309

QY 290 KPPRDHVFOINNFEALKTIONQREKIFAIBGTQGTSSSFHEMSQEGFAAITSNGPL 349
Db 310 EPKLNNTFYIQEYNGLGLDLNQLKKIYNIEGSKAHRDRQKELSQSGFSVYQESVI 369

QY 350 LSTVGSYDWAGVFLYT---SKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVLG 406
Db 370 VGSVGSNDWRGALYEVMSGSKPRQTEITDPAVN---KDSYMGYSTVLGMRHGVSLTFSG 426

QY 407 APRYOHIGLVAMFRQNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHY 466
Db 427 APRAEHTGLTTLTKNESTWTVNRNINGEQISYFGASLUSLDDVDSGDSFLLVGAPLP 486

QY 467 YE-QTRG-QQVSVCPPLPRQARWQCDVLYGEQ--GQPWGRFGAALTVLGDVNGDKLTD 522
Db 487 YQSQPRAEGRLYVYTL-----SEYQSKTLQSTTGRFATSLASLKDNLGDLSD 535

QY 523 VAIGAPGEEDNRGAVLPHGTSGSGISPSHS--QRTAGSKLSRQLYFGQSLSGQDLTMD 581
Db 536 VAVGAPLE--NEGVMYIYLDGTHGINPEHAPQIPARSVLPGLOQFGVLSGQDMND 593

Qy	582	GLVDLTGVAQGHVLLRSQVLRVKAIMEFNPREVARNVPECNDQVVKGKEAGEVRVCLH	641
Db	594	NLPDVIQTGGIVLLNARVMSVSAQLSPNPMELISINYPECPGS--NAPNAFNLTSCTF	651
Qy	642	VQKSTRDLREGQIQSV--VTYDLALDSGRPHSRVAFNFKNSTR--RQOVVLGLTQTCE	697
Db	652	VTERTGS--TGSLEKKLVLSNLNLDVWVGMSRGFPDPMDSRSRTLQOSVLLDSSGCS	708
Qy	698	TLKLQLPNCIEDPVSPVILRLNFS---LVGTPLSAFNGLRPLVLAEDAQRLELTALPFPEK	753
Db	709	NFSIFMLRCVADIVSPCLKIRNMFSTQMLSGNSLAVL-DIQSRTEEVEVL-----FQR	761
Qy	754	NCGNNDNICDDLSITFSFMSLDCLLVGGPREFNVTYVRNDEGDSYRTQVTFPPFLDLSY	813
Db	762	NC-NSNSCVADLKLNFSTN-DTLVVENQAHTVLVSLANPGDSDSYNTSVILHYPEGLSL	819
Qy	814	RKYSTLQVQORSQRWRLACBSASTEVSGALKTSCTSIINHPIPFENSEVTFNITFDV---	870
Db	820	SKFDAIKPSRTR-----SSCDDRDSGATNRITCTSIPLPVYRSRGTTTQFLGTFRVTKW	871
Qy	871	DSKASLGNKLLKANVTSENNMPRTNTEFOLELPVKYAYVMVVTSHGV-STKYLNFITAS	929
Db	872	DYDWSDRMWTITANSNNGM---SDMSVRSIPQVFAVELAISLVAEDSVTYLNFSL	928
Qy	930	ENTSRVMOHQVQVSNLQORSUPISLVLVPLVRLNQTVIWDRPQVTFSENLSSTCHKERL	989
Db	929	DRGPKPLNIYKVVNLGKGLPVSVTLSPCQ-----TTHVTLTPHNFSMQ	974
Qy	990	PSHSDFLAELRKAPVNC-----SIAVCQRIQCDIPFFGLOE	1027
Db	975	EVHSHSFSSYHQ--IIMCLLNKHLFFSPELSAVQTRTTGRSLWC-----VSSISGTGEI	1026
Qy	1028	FNATLKGNL-----SFDWYIKTSHNHLILVSTAELFNDSVFTLLPG-QGA	1072
Db	1027	FRSSV--NLMAEAVLQNVKEYESKYSFY-BFRRDHVFNIS-AELNFNNTSYNOSTGLKYN	1082
Qy	1073	FVRSQTEKVEPFVPPNPPLPIVGVSSVGGLLLLALITAALYKLGFFKRYQKD	1124
Db	1083	PHRSQTEKVEFVIPPSEMLIVCTGAVGGFFFLIIILLLLKCGFFKERNPD	1134
RESULT 12			
Q8HZV0			
ID	Q8HZV0	PRELIMINARY;	PRT; 927 AA.
AC	Q8HZV0;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 (Fragment).		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI TaxID=9913;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;		
RT	"Sequence of the alpha subunit of bovine lymphocyte function-		
RT	associated antigen 1."		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF440778; AA63636.1; --		
FT	NON TER	1	1
FT	NON TER	927	927
SQ	SEQUENCE		927 AA; 102523 MW; 02E3CF09917081EC CRC64;

Qy	234	FKILFLLTDEGKPGDPLGVEDV	PELDRBGVIRYVIGVCDAPRSEKSRQELNTVASKPPR	293
Db	64	TKVLIITDGK---PPTNTLMRPKTS	---RSLGIGKNFKTKESQALHQFASKPVE	116
Qy	294	DHYFQNNFEALKTIONLREKIF	ATEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTV	353
Db	117	EFVKIILDTTEKLDLTFELQK	YIVIEGTSKQDLTSFNMELSSGISADLSGEGHVGWGVN	176
Qy	354	GSYDWAAGGVF-LYTSKEKSTF	FINMTRVDSMDMDAYLGYAAA-IILNRNVOSILVLAGPARYQ	411
Db	177	GAKDWAGGFLLDKADLKSSTF	VGNEBPLTVESRAGVLGYTVTRLPGRGTMSLLATGAPKYQ	236
Qy	412	HIGELVAMPRO--NTGHWESNAV	KGTQICGAYFGASLCSDVDNSNGSTDLVLGAPHYYEQ	469
Db	237	HVGRVLLFOQPKRGPGWPSQI	QIEDGIGSYFGGELCGVDVDRDGETELLAAALPYGE	296
Qy	470	TRGQGVSVCLPRGQARWQCDA	VLGEGQQWGRFGAALTVLGDVNGDKLTDVAIGAPG	529
Db	297	QRGRVFIY---OKIQLEFQM	YSELQETGYPLGRFGAAIAALTDINGDELTDVAVGAPL	353
Qy	530	EEDNRGAVLYFHTGTSGGIS	PSHSQRIAGSKSLPRLOYFGQSLSGGQDLTMDGLVDLTVG	589
Db	354	EE--QGAVYIFNGQOG-GLSP	RSQRIEGTQMFSGIOWFGSRSHGVKDLGGDGLADVAVG	410
Qy	590	AQGHVLLLRSPVLRVKAIME	FNPREVARNVFECNDQVVKGEAG-EVZVCLHVQK--ST	546
Db	411	AEGQVILSRPVDIITISVSF	SPASIPVHEHVESYSTSNQKKEGVNLTVCFOVKSLLJST	470
Qy	647	RDLREGIOQSVVTDLALDS	GRPHSRVAFNETKSTRQTQVGLGTQTCETLKLQLPNC	706
Db	471	----FQGHVANIUTYTLQD	CHTRSRGILFPGGKHKLIGNTAVTPV-KSCFVFWPHFPIC	525
Qy	707	IEDVPSPVLRLNPSL---VGT	PLS--AFGNLRPLVAEDAQRFTALFPPEKNCNGNDNIC	761
Db	526	IQDLISPIVSLSVSLWE	EBEGTPRDPRALDRDIPPIKPSPHLETKEIPFEKNCGEDKNC	585
Qy	762	QDLSITFSFMSLDCLVGG	PREFNTVTVVRNDEDSYRTQVTFPPFLDLSYRKVSTLQN	821
Db	586	EADLKAFSDMRKILRLT	PSASLVRLTALRNDAEAYVQVTLSPFQGLSPFRKVEIL--	643
Qy	822	QRSQRWLACESASSTSE	YSGALKSTSCSINHPIFPENSEVTENIFDVSOKASLGNKLL	881
Db	644	-KPHSHVPVGCBELPEE	AVVHS--RALSCNVSSSIFGEDSMVDIQVNFNTLQKSGWDPIE	701
Qy	882	LKANVTS-----ENNMP	RTNKTEFQELPVPKYAVYVVTVSHGYSTKYLNFPTASENTS	933
Db	702	LQANVSCNNEDSLL	EDNSATTS-----IPVMYPIVNLTKDOENSTLYISFTKPSPRI	754
Qy	934	RVMOHQVQV-----SNL	QORSPLSLVFLVPVRLNQTVI---WD---RPQVTFSS-ENLSST	982
Db	755	HVHKHIYQVRIQPSY	NDMP-PLEALVRVPRVHSEGLITHKNSIQMEPPVNCSPRNLESP	813
Qy	983	CHTKERLPSHSDFLAEL	KAPVNCSTAVCQRIQCDDIPFGEGIEEFNATLKGNSLFDWYI	1042
Db	814	SDEAE-----SCSFGT	--EFRCPIDF---RQELVQVNGVWEILRGTI	850
Qy	1043	KTSHNHLIIVSTAEIL	FNDSVFTLLPQGGAFVRSQOTETKEVPEFVNPLPLIYGVSSVGG	1102
Db	851	KAS-SMLSSCLSLAIS	SNSSKHFLHLHGRNASM-AQVVMKVLDLYVEKEMLYLYVLSGIGGL	908
Qy	1103	LLLALITAALYKLG	FFKFR	1120
Db	909	LLLELITIALYKVG	FFKFR	926

RESULT 13	
OB8340	
ID	OB8340
AC	PRELIMINARY; PRT; 1167 AA.
DT	01-NOV-1998 (TRENBLrel. 08, Created)
DT	01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)


```

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02234; VWF; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.6%; Score 1154.5; DB 11; Length 1167;
Best Local Similarity 28.9%; Pred. No. 1e-76;
Matches 356; Conservative 214; Mismatches 458; Indels 205; Gaps 43;

QY 1 FNLDTEA--MTFQENARGQSVQLQGSVVVGAPQEITVAANRQS-----LYQCDY 52
DB 20 FNNVDWAWYALQCAPVLSLLHLDPSN-----NOTCLLVARRSSNNTAALYRCAL 74
QY 53 STGSCPIRLQVPEAVNMSLGSLAATT--SPQLLAC-GPTVHOTCSENTVYKGLCLF 109
DB 75 SI-SPOEIAQ-PVEHICMEKRGYQGVTLVGNHNGVLVCIQVQARKFRSLNSELTCACSL 132
QY 110 FGSNLQAOQKPFPEARLG-----C-----PQE 131
DB 133 LTNLDLQAOYPSDLLEGFLDGAHVDSGYCRSKSGSTGEEKSARRRTVEEED 192
QY 132 DSDFIADLGGSGIIPHDFRRMKEMVSTVMEQL--KSKTLFSLMOYSEEFRIHFTKEF 189
DB 193 GTETAIVLDSGSIERSDFQAKNFISTMWRNFEKCFECNFALVOYGAIVQTEFDLQES 252
QY 190 QNNPNRSLIKPTQLLGRTHRTATGLRKVVRELNITNGARKNAFKILFLLTGDGKFGDP 249
DB 253 RDINASLAKVQSVQVKEVTKTASAMQHVLDNIFPSRGSRKALKVMVVLTDGDFGDP 312
QY 250 LGYEDVIPELDRGVIRYVGVGDAPRSEKSRQELNTVASKPPRHVFOINNFEALKTIQ 309
DB 313 LNTTTVINSFKMGQVRFVGVGDGAFKNNNTYBELKLIASDDPKAEHTFKVTNYSLDGLL 372
QY 310 NQUREKIFATEGTQSGSSFEHMSGQESAAITSNGP--LLSTVGSYDWAGGVFLY-TS 367
DB 373 SKLQQRIVHMEGT---VGDALQVLAQTGFSAQLDKGVLLTGVAFNWSGGALLYSTQ 429
QY 368 KEKSTTINMT-RVDS-DMNDAYLGAAAIILNRVQSLVLGAPRYOHIGLVAMFRQWTGM 425
DB 430 NGRGCFNLQAKEDSRVTQVSYLGSVLAVLHKAHGISYVAGAPRHKLARGAVELRKEDR- 488
QY 426 WESNA---NVKGTQICAYFGASLCSDVDNSNGSTDLVLGAPHYEQTGGQGVSCPLPR 482
DB 489 -EDAFVRRIEGQMGSYFSGVLCPDVIDMDGTTDFLLVAAPFYHIRGEGRVYVQVPE 547
QY 483 GQARWQCDVAVLGEQGPNGRGCAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFLSLHTLSGHFGLTNSRFGFMAAVGDIINQDKFTDVAIGAPLEGFAGDGASYS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKPLRQYFGQSLSGGQDLTMDGLDLTVGAQGHVLL 596

```

```

DB 607 VVIYNGHSG-GLYDPSQOIRASSVGLHYFGMSVSGGLDFNGDGLADITVGSRDSAV 665
QY 597 LRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSWTETP-----DALPMVFIGM--DVNLCFEVDSSVWASEPGLREM 715
QY 654 QIOSVVTYDLALDLSGRPHSRVFNETKNSRTRQTVLGLTQTC----- 696
DB 716 FLNFTVDV-----TKQRRLQCEDSSGCQCLRWKNGSGFLCEHFWLI 760
QY 697 ETLKQLPNCIEDPVSPIVLRNFSLVGTFGLSFAFNLR-----PVLAEQAQLFTALF--P 750
DB 761 STEEL-----CEEDCFSNITIKVTYE-----FOTSGRRDYPNPTL--DHYKEPSAIFQLP 809
QY 751 FEKNGCNDNICODDLISITFSFMSLDCLVVGSGPREFNVTVTVRNDGEDSVYRTOVTFEPLD 810
DB 810 YEKDCNKVFCIAEIQLTN-ISQELVVGVTKEVTMNIISLTNSGEDSVTMTNMLNYPN 868
QY 811 LSYRKVSTLQONORSQSRWLACESASSTEVSGALKSTSCSINHPIPPENSEVTNITFDV 870
DB 869 LQFKKI-----QKPVSPDVQCDPKPV---ASVLVMNCKIGHPIIL-KRSSVNVSVTWQL 918
QY 871 DSKASLGNKLLKANVTSENMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASE 930
DB 919 EESVFPNRTADITVTISNSNEKSLARETR---SLOFRHAFIAVLRS--PSVVMYN--TSQ 971
QY 931 NTSRVQMHOYQVSNLQORSPLISLVFLVPRLNQTVINDRPOVTPSENLSST-----CHT 985
DB 972 SFSDHKEFFNVHGENLFGAVFQLCVPIKLODF-----QIVRVKNTKTQDHTECTQ 1025
QY 986 KERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEBFNATLKNLSFDWYIKTS 1045
DB 1026 SQEPACGSDPVQHVKEHWSVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILPNDVSVTLLPQCGAFVRSQTETKVEPF-----EVPNPLPLIV 1095
DB 1066 HTKQLLRDVELSELPIGELISFNKSLYEGLNAE-----NHRTKITVIFLKEEETRSLPLII 1119
QY 1096 GSSVGGLLLLALITAALYKLGFFKQYKDMNSE 1128
DB 1120 GSSIGLLVLVVIILFKCGFFKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

```

DR SMART; SM00191; Int. alpha; 4.
DR SMART; SM00327; VWA_1
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON TER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;

Query Match 18.8%; Score 1109; DB 11; Length 1167;
Best Local Similarity 28.0%; Pred. No. 25e-73;
Matches 332; Conservative 206; Mismatches 424; Indels 244; Gaps 37;

QY 47 LYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPQLLACQPTVH----- 94
DB 67 LYQC-----ATSIIPDEIGCPVEHILMPKGRYQGV 98
QY 95 -----QTCSENTYVKGCLFGLFSGNLROQKPEALRG----- 127
DB 99 LVRNHNGLVLCIQVQRKPRSLNSELTGACSLTLPNLDLOAQVPSDLEGVLDLGASVNS 158
QY 128 -----CPOED-----SDIAFLIDGSGSIIPHDFRRMKEW 157
DB 159 GDYGSKGSGTEETKSARRLRQAVEDEEAEETEAIIVLDGSGSIEPSDFQAKADFI 218
QY 158 STVMEQL--KKSRTLFLMQLYSEEPRIHFTFEFQNNPNPRSLIKPITQLLGRTHRTA 215
DB 219 STMMNFYKCFECNFALVQGVVQTETFDLLDSRDINASLAKVQSIQVQKVTKTASAM 278
QY 216 RKVRELNIITNGARKNAFKILFLTDGKFGDPLGYEDVIPELDREGVIRYVIGVDAP 275
DB 279 QHVLNFIPIPSGRSKKALKVMVLTDGDIIFRDPLNLTITVSSSKMQGVVRFVIGVNAF 338
QY 276 RSEKROELNVTASPRDHVFOINNFALKTIONLREKIPALGETGTGSSSFEHMS 335
DB 339 ENNTYREKLIASDPKAAHPTKVTNYSALDGLLSKLQRIHMEGT---VGDITLQYOLA 395
QY 336 QEGFSAAITNSGP--LLSTVGSYDAGGVFLY--TSKEKSTFTINMTVDSDMND 390
DB 396 QTGFSAQILDKQVLLGTVGAFNWSGGALLYNTQNGRFLNQT--AKEDFRAAQVSYLGY 454
QY 391 AAAIILNRVQSLVIGARYQHIGLVAMPFRONTGWESNANVKGTOIGAYGASLCSVDV 450
DB 455 SVAALHKAHGVSYVAGAPRHKLKRGAVFELQKEDGETFMRRIEGQMGSYFGSVLCPVDI 514
QY 451 DMSGSTDLVLICAPHYECTRGQVSVCLPRGRARWOCDAVLVGEQGWGREGAALT 510
DB 515 NMDGITDFLLVAAPFYHIRGEGRVYVTRV--HEQDAPFSLVTLTSGYGLTSSRRCFAMA 573
QY 511 VLGDVNGDKLTDVAICAP-----GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR 564
DB 574 AVGDIHQDKFTDVAIGAPLEGFAGDGASVGSVYIYNGHSG--GLHASPSQIRASSVALG 632
QY 565 LOYFGSLSGGQDLTWDGLVLTGAQGHVLLRSQPVLRVKAIWFEFNPREVARNVFEEN 624
DB 633 LYYFGMSVSGGLDFSGDDLADITVGSQDVAVVLRSPVVDLTVSMFTPTDALP----- 685
QY 625 DQVVKGEAGEYRVCLHWKQK---TRDLREGQIQSVVYTDLALDSGRPHSRVAFNETKN 681
DB 686 ---MAFKQKNDVELCKVDSSAVPSEPLRGMSLFTVDVDV-----TQ 727
QY 682 STRRQ-----TQVLGLTQTCTETLKLQPNCEIDPVSPIVLRNFS 721
DB 728 KQRLQCADRSCCQSLMKWSGSSSLCEHFLISTEEL-----CEDDCFSNITIKVSYE 781
QY 722 LVGTPLSAPGNLRPVLAEDAQLFTALP--PREKNGNDNIQDDLSITFFSPMSLDCLVV 779
DB 782 F--OTSEERNHNPNIL--DHYKEPSAIFQLPYEKDCKNKVFCIABIQLTTLAISQOD-LVW 837
QY 780 GGPRENVTYVRNDEGDSYRQVTFFFPLDLISYKVKVSTLQNRORSORSLWACASSTE 839
DB 838 GITKEVTMNLISLNSGEDSYMTNMLNYPNNOFKKI-----QKPLSPDIQCDPKPV- 890
QY 840 VSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTE 899

891 --ASVLVNMCKIGHPII-KRSSVNVSVTWOLEESSIFPNRTADITVTISNSNEKSLARETH 947
900 FOELPVKYAVYVWVTSYKYNFTASENTSRVQHQYQVSNLQSRSLPISLFLVLP 959
948 ---SLQFRHAFIAVLSR--PSVVMYMN--TSOSSSDHKKEFPFNVHGENHFGAVFQLQICVP 1000
960 VRLNQTVWDRPQVTFSENLSST-CHTKERLPSHSDFLAELRKAPVNVVNCISVACQIQCD 1018
1001 ITRDL-----QIIRVKHLTKTQAHT-----ECTQSQEPTCGSDPVQNV--- 1039
QY 1019 IPFPGQEEFNATL-----KGNLSFDWYIKTSHNHLII--VSTAEIL-----FNDSVFTL 1066
DB 1040 -----EWHSVICAITSNKENVTVAEISMGHTKQLDRDISLQILGSIENKSLIEG 1092
QY 1067 LPQGAFVRSQETKVEPF-----EVNPNPLPIVSGSVGGLLLLALITAALYKLGFFKQY 1122
DB 1093 LNAE-----NHRKTIVIFLKEEKPHSLPLIIGSSIGLLVLLVIAILFKCGFFKRY 1146
QY 1123 KQMMSE 1128
DB 1147 KQNLSE 1152

RESULT 15
O42094
ID O42094 PRELIMINARY; PRT; 1171 AA.
AC O42094;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ALPHAI integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
the alpha1 integrin gene";
RL J. Biol. Chem. 272:26643-26651 (1997).
DR EMBL; AB000470; BAA23160.1; -;
DR EMBL; AB000471; BAA23161.1; -;
DR HSSP; P17301; IAOX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5;
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int. alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Query Match 17.7%; Score 1041; DB 13; Length 1171;
Best Local Similarity 28.0%; Pred. No. 3e-68;
Matches 340; Conservative 213; Mismatches 511; Indels 152; Gaps 46;

QY 2 NLDTENAMTFOENARG-FGQSVVQL---QSGRVVVGAPQETVAANQGRSLYQCDYSTGSC 57
DB 19 NVDVKNAMTFSGPLEDMFGYTVQYENEEGKWLIGSLVGVQPEKRTGDVYKCPVGRDSQ 78
QY 58 EP-IRLQVP-----VEAV--NMSLGLSLAATTSPQPLLACGPTVHQTCSNTYVKGCL 107
DB 79 SPCKLNLPDATSVPNVMEVKENMTLGTTL-VTPNKGGLACGPLYAYKCGRLHYTTGVC 137

GenCore version 5.1.6
Copyright (c) 1993 - 2003: Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481b-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPGGAPFP 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries:

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852	99.5	1153	19	Human Beta-integrin
2	5852	99.5	1153	21	Human CD11b protei
3	5852	99.5	1153	23	Human Beta2 integr
4	5852	99.5	1153	23	Human integrin 1 a
5	5852	99.5	1153	23	Integrin Mac-1 alp
6	5842	99.3	1153	11	Alpha subunit of M
7	3450	58.6	1163	11	p150.95 alpha subu
8	3436	58.4	1163	19	Human Beta-integrin
9	3436	58.4	1163	21	Human CD11c protei

10	3436	58.4	1163	23	ABG61470	Human Beta2 integr
11	3434	58.4	1163	24	ABU07406	Protein differenti
12	3401	57.8	1161	18	AAW78166	Human beta-2 integ
13	3401	57.8	1161	18	AAW23049	Human beta 2 integ
14	3401	57.8	1161	19	AAW72825	Human alpha-d. Ho
15	3401	57.8	1161	19	AAW65089	Human Beta-integrin
16	3401	57.8	1161	19	AAW57491	Human Beta2 integr
17	3401	57.8	1161	20	AAW73342	Human alpha d clon
18	3401	57.8	1161	21	AAW07359	Human alpha d prote
19	3401	57.8	1161	23	ABG61468	Human Beta2 integr
20	3385.5	57.5	1161	18	AAW23064	Human beta 2 integ
21	3385.5	57.5	1161	19	AAW72837	Human alpha-d deri
22	3385.5	57.5	1161	19	AAW65106	Human Beta-integrin
23	3385.5	57.5	1161	20	AAW73343	Human alpha d prote
24	3385.5	57.5	1161	21	AAW07376	Human alpha d integ
25	3385.5	57.5	1161	23	ABG61485	Human Beta2 integr
26	3226.5	54.8	1161	16	AAW78169	Rat alpha-d subuni
27	3224.5	54.8	1161	18	AAW23062	Rat beta 2 integr
28	3224.5	54.8	1161	19	AAW72824	Rat alpha-d #1. R
29	3224.5	54.8	1161	19	AAW60004	Rat alpha d poly
30	3224.5	54.8	1161	21	AAW07374	Rat alpha d protei
31	3224.5	54.8	1161	23	ABG61483	Rat Beta2 integrin
32	3217.5	54.7	1161	19	AAW65104	Rat beta-integrin
33	3217.5	54.7	1161	20	AAW73345	Rat alpha d integr
34	3212	54.6	1161	20	AAW23061	Mouse alpha d #2
35	3212	54.6	1161	19	AAW72836	Mouse beta-integrin
36	3212	54.6	1161	19	AAW65103	Mouse alpha d poly
37	3212	54.6	1161	19	AAW60003	Mouse alpha d integ
38	3212	54.6	1161	20	AAW73347	Mouse alpha d prote
39	3212	54.6	1161	21	AAW07373	Mouse alpha d integ
40	3212	54.6	1161	23	ABG61482	Mouse alpha-d subu
41	3208	54.5	1161	16	AAW78168	Rat beta 2 integr
42	3203.5	54.4	1151	18	AAW23059	Rat alpha-d #2. R
43	3203.5	54.4	1151	19	AAW72834	Rat beta-integrin
44	3203.5	54.4	1151	19	AAW65101	Rat alpha d poly
45	3203.5	54.4	1151	19	AAW60001	Rat alpha d poly

ALIGNMENTS

RESULT 1

AAW65090

ID AAW65090 standard; Protein; 1153 AA.

XX AAW65090;

XX AC

XX DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

XX DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;

XX KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;

XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;

XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX PN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-0485618.

XX PR 07-JUN-1995; 95US-0485618.

XX PR 23-DEC-1993; 93US-0173497.

XX PR 05-AUG-1994; 94US-0286889.

XX PR 21-DEC-1994; 94US-0362652.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van DER VIEREN M;

XX XX

DR WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-I diabetes

XX

PS Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 19; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDENAMTQENARGGQSVQVQGGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

DB 17 FNLDENAMTQENARGGQSVQVQGGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFLFGSLNRQOPQK 120

DB 77 RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFLFGSLNRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKXKTLPSLMQYSEEP 180

DB 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKXKTLPSLMQYSEEP 196

QY 191 RIHFTFEFQNNPRSLKPIITOLLGRTHATGLRKVVRELFTNGARKNAFKILFL 240

DB 197 RIHFTFEFQNNPRSLKPIITOLLGRTHATGLRKVVRELFTNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 300

DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 316

QY 301 NFEALKTIONQLREKIFAEIGTQSGSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360

DB 317 NFEALKTIONQLREKIFAEIGTQSGSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376

QY 361 GVFLYTSKESKSTFNTMTAVDSMDNDAYLGXAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420

DB 377 GVFLYTSKESKSTFNTMTAVDSMDNDAYLGXAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTOICAYFGASLCSDVDNSGSTDVLVIGAPHYEQTRGGQVVCPL 480

DB 437 QNTGWESNANVKGTOICAYFGASLCSDVDNSGSTDVLVIGAPHYEQTRGGQVVCPL 496

QY 481 PRGQARWQCDVAVLYGEGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540

DB 497 PRGQARWQCDVAVLYGEGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556

QY 541 HCTSGSGISPSHSQRIAGSKLSPRLOYFGQSISGGQDLTMQGLVDLTVGAGQHVLLRSQ 600

DB 557 HCTSGSGISPSHSQRIAGSKLSPRLOYFGQSISGGQDLTMQGLVDLTVGAGQHVLLRSQ 616

QY 601 PVLVRKATMEFNPREVAENVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

DB 617 PVLVRKATMEFNPREVAENVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHRAVFNETKNSRRTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 720

DB 677 YDLALDSGRPHRAVFNETKNSRRTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFQNLRPVLAEDAQRFLTALFPFENKNCNDNIQDDLSITTSFMSLDCLVWG 780

DB 737 SLVGTPLSAFQNLRPVLAEDAQRFLTALFPFENKNCNDNIQDDLSITTSFMSLDCLVWG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840

DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIEPENSEVTNITFDVDSKASLGNKLLKANKVTSNNPRTNKTFF 900

DB 857 SGALKSTSCSINHPIEPENSEVTNITFDVDSKASLGNKLLKANKVTSNNPRTNKTFF 916

QY 901 QLELPVKYAVYVMVTVTSHGVSTKYLNTFASNTSRVVMQHVQVSNLQQRSLPISLVLPVP 960

DB 917 QLELPVKYAVYVMVTVTSHGVSTKYLNTFASNTSRVVMQHVQVSNLQQRSLPISLVLPVP 976

QY 961 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020

DB 977 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036

QY 1021 FFGIQEENATLKGNSLFDWYIKTSHNHLIIUSTAILFNDVSFTLLPGQAFVRSQTET 1080

DB 1037 FFGIQEENATLKGNSLFDWYIKTSHNHLIIUSTAILFNDVSFTLLPGQAFVRSQTET 1096

QY 1081 KVEPFEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137

DB 1097 KVEPFEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX AAB07360;

XX 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

XX Human; macrophage infiltration inhibition; alpha_d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

OS

XX WO200029446-A1.

PN

XX 25-MAY-2000.

PD

XX 16-NOV-1999; 99WO-US27139.

PF

XX 16-NOV-1998; 98US-0193043.

PR

XX 08-JUL-1999; 99US-0350259.

XX (ICOS-) ICOS CORP.

PA

XX Gallatin MW, Van Der Vieren M;

PI

XX WPI; 2000-387751/33.

DR

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites

PT

XX Example 5; Fig 1; 270pp; English.

PS

XX Integrins are a class of membrane-associated molecules that participate

CC

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha_d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha_d. The alpha_d gene and protein may be useful in therapy for diseases linked to aberrant alpha_d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (IAD). In addition, anti-alpha_d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

XX Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 21; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQRSLYQCYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQRSLYQCYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOOK 136
QY 121 FPEARLGCQEDSDIAFLDGSIIIPHDFRMKEMVSTVMEQLKKSKTLFSLMQYSEEF 180
DB 137 FPEARLGCQEDSDIAFLDGSIIIPHDFRMKEMVSTVMEQLKKSKTLFSLMQYSEEF 196
QY 181 RHFTFKFQNNPNRSLIKPTOLLGRTHATGIRKVVRELFNITNGARKNAKILVLI 240
DB 197 RHFTFKFQNNPNRSLIKPTOLLGRTHATGIRKVVRELFNITNGARKNAKILVLI 256
QY 241 TQGEKFGDPLGVEDVPELDREGVRYVIGVDGAFRSEKSRQELNATVASKPRDHVFQIN 300
DB 257 TQGEKFGDPLGVEDVPELDREGVRYVIGVDGAFRSEKSRQELNATVASKPRDHVFQIN 316
QY 301 NFEALKTIQNLREKIFALEGQTQSSSFEHEMSOEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFALEGQTQSSSFEHEMSOEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSCPL 480
DB 437 QNTGMWESNANVKGTQIGAFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSCPL 496
QY 481 PRGQARMQCDVLYGEOQPGFRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLF 540
DB 497 PRGQARMQCDVLYGEOQPGFRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFQCSLGGQDLTMDGLVDLTGVAQGHVLLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFQCSLGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616
QY 601 PVLVRKATMEFNPREVARNVFCNDQVKGKEGVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLVRKATMEFNPREVARNVFCNDQVKGKEGVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGTTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSRROTQVLGTTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPFCKKNCNDNICQDDLSITFSFMSLCLVVG 780

DB 737 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPFCKKNCNDNICQDDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFPPFDLSYKSVSTLQORSORSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRQVTFPPFDLSYKSVSTLQORSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSPLVLPV 960
DB 917 QLELPVKYAVVMVTSHGVSSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSPLVLPV 976
QY 961 RLNQTIVIDRQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVVCISIAVCORIQCDIP 1020
DB 977 RLNQTIVIDRQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVVCISIAVCORIQCDIP 1036
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSQTET 1080
DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSQTET 1096
QY 1081 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137
DB 1097 KVEPPEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX
AC ABG61469;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human Beta2 integrin alphaCD11b subunit.
XX
KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS Homo sapiens.
XX
PN WO200230980-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32059.
XX
PR 13-OCT-2000; 2000US-0688307.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin WM, Van Der Vieren M;
XX
DR WPI; 2002-463260/49.
XX
PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX
PS Example 5; Page 191-194; 270pp; English.
XX
CC The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
 CC spinal cord injury victim. The method also involves the use of
 CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
 CC molecule, vascular cell adhesion molecule). The method is useful for
 CC promoting locomotor recovery, inhibiting locomotor damage, limiting
 CC locomotor impairment, or limiting autonomic and sensory dysfunction
 CC following spinal cord injury. In particular, the spinal cord injury
 CC comprises compression of the spinal cord. The antibodies are also useful
 CC for reducing inflammation at the site of a central nervous system injury.
 CC The specification also details the identification of Beta2 integrin
 CC alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2
 CC integrins are implicated in diseases such as LAD (leukocyte adhesion
 CC deficiency, inflammatory response, diabetes, multiple sclerosis,
 CC arthritis, graft atherosclerosis, inflammatory bowel disease,
 CC Crohn's disease, ulcerative colitis, immune complex alveolitis
 CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
 CC sequence included for comparison with the Beta2 integrin alpha protein
 CC sequences.

XX SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 23; Length 1153;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFQGSVQLOQSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
 Db 17 FNLDTENAMTQENARGFQGSVQLOQSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATSPQQLACGTVHTQTCSENTYVYKGLCFLFGSLNRQQPQK 120
 Db 77 RLQVPVEAVNMSLGLSLAATSPQQLACGTVHTQTCSENTYVYKGLCFLFGSLNRQQPQK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKKEVSTVMEQKKSKTLFSLMOYSEEF 180
 Db 137 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKKEVSTVMEQKKSKTLFSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPRSLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILFL 240
 Db 197 RIHFTFKFQNNPRSLVKIPITQLLGRTHATGLRKVRELFINITNGARKNAFKILV 256

Qy 241 TDGEKFGDPLGVEDVPELDEGVIRYVIGVDAPFRSEKSKQELNTVASKPPRHVQFIN 300
 Db 257 TDGEKFGDPLGVEDVPELDEGVIRYVIGVDAPFRSEKSKQELNTVASKPPRHVQFIN 316

Qy 301 NFEALKTIONOLREKIFAIECTOTGSSSSFEHMSQEGFSAITSNGLLSTVGSYDWAG 360
 Db 317 NFEALKTIONOLREKIFAIECTOTGSSSSFEHMSQEGFSAITSNGLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 Db 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
 Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496

Qy 481 PRGQARWQCDVLYGEGQPNRGAAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540
 Db 497 PRGQARWQCDVLYGEGQPNRGAAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616

Qy 601 PVLRVKAIMERNPREVAENVFECNDQVYVKGAEGRVCLHVQKSTRDLREGQIQSVVT 660
 Db 617 PVLRVKAIMERNPREVAENVFECNDQVYVKGAEGRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNFTKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
 Db 677 YDLALDSGRPHSRVAFNFTKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736

Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFFKKNCGNDNICQDDLSTTFSMSLDCLVVG 780
 Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFFKKNCGNDNICQDDLSTTFSMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSVRKYSTLQNRQSRQSWRLACESASSTEV 840
 Db 797 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSVRKYSTLQNRQSRQSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
 Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTFASENTSRVVMQHQYQVSNLQORSPLISLVLVPV 960
 Db 917 QLELPVKYAVYVMTSHGVSTKYLNTFASENTSRVVMQHQYQVSNLQORSPLISLVLVPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
 Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036

Qy 1021 FFGIQBEFNATLKGNSLSPDWYIKTSHNHLILVSTABILFNDISVFTLLPGQCAFVRSOTET 1080
 Db 1037 FFGIQBEFNATLKGNSLSPDWYIKTSHNHLILVSTABILFNDISVFTLLPGQCAFVRSOTET 1096

Qy 1081 KVBPFFVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
 Db 1097 KVBPFFVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 4
 AAU80252
 ID AAU80252 standard; Protein; 1153 AA.
 XX AC AAU80252;
 XX DT 15-JUL-2002 (first entry)
 XX DE Human integrin 1 alpha-M subunit protein.
 XX KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
 KW inflammatory disease; autoimmune disorder; Crohn's disease;
 KW human immunodeficiency virus; HIV; myocardial infarction;
 XX KW Sjorgen's syndrome; rheumatoid arthritis.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Misc-difference 499..500
 FT /note= "Encoded by GGG CAG AGG"
 XX WO200218583-A2.
 XX PD 07-MAR-2002.
 XX PF 31-AUG-2001; 2001WO-US27227.
 XX PR 01-SEP-2000; 2000US-229700P.
 XX PA (BLOO-) CENT BLOOD RES INC.
 XX PI Springer TA, Shimoaka M, Lu C;
 XX WPI: 2002-382964/41.
 DR N-PSDB; ABK50046.
 XX Modified integrin-I or integrin I-like domain polypeptide useful as an
 PT immunogen to produce antibodies specific to polypeptide, comprises a
 PT disulfide bond such that polypeptide is stabilized in a desired
 PT conformation -
 XX Disclosure; Page 109-112; 112pp; English.
 XX This invention relates to a modified integrin-I or integrin I-like
 CC

domain polypeptide comprising at least one disulfide bond so that the domain is stabilised in a desired conformation. The polypeptide of the invention may have anti-inflammatory or immunosuppressive activities. The polypeptides of the invention have an open conformation and are useful as immunogens to produce antibodies that selectively bind to integrin I-domain; and for identifying a modulator of integrin activity, or of interaction of an integrin and a cognate ligand. The polypeptide of the invention, or antibodies (preferably anti-LFA-1 antibody) is useful for treating or preventing an integrin mediated disorder which is an inflammatory or autoimmune disorder in a subject and for inhibiting the binding of an integrin to a cognate ligand such as Crohn's disease, nephritis, human immunodeficiency virus (HIV), myocardial infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic composition comprising the peptide of the invention is useful for treating an integrin mediated disorder in a subject. The polypeptides and/or active or antigenic fragments are useful as reagents for diagnosis of integrin-mediated disorders. The present sequence represents the human integrin-1 alpha-M protein subunit used to generate the mutant polypeptides of the invention.

XX SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 23; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLYQCDYSTGSCBPI	60
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLYQCDYSTGSCBPI	76
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFLFGSLNRQQPOK	120
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFLFGSLNRQQPOK	136
QY	121	FPEALRGCEQSDSDAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSBEF	180
DB	137	FPEALRGCEQSDSDAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSBEF	196
QY	181	RIHFTPFKEQNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFNTNGARKNAKILFL	240
DB	197	RIHFTPFKEQNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFNTNGARKNAKILVVI	256
QY	241	TDGEXFGDPLGVEDVIPLEDRGVIRYIVGVGDGAFRSEKSRQELNTVASKPRDHVFQIN	300
DB	257	TDGEXFGDPLGVEDVIPLEDRGVIRYIVGVGDGAFRSEKSRQELNTVASKPRDHVFQIN	316
QY	301	NFEALKTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPLSTVGSYDWAG	360
DB	317	NFEALKTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLCAPRYQHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLCAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVKGQTIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYEQTRGGQSVVCP	480
DB	437	QNTGMWESNANVKGQTIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYEQTRGGQSVVCP	496
QY	481	PRGQARWQCDVAVLGEQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
DB	497	PRGQARWQCDVAVLGEQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	556
QY	541	HGTSGSGI:SPSHQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ	600
DB	557	HGTSGSGI:SPSHQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ	616
QY	601	PVLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
DB	617	PVLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSRAVFNETHKSTRQTQVLGTLTQTCETLKLQLPNCIEDPVSPIVLRNF	720
DB	677	YDLALDSGRPHSRAVFNETHKSTRQTQVLGTLTQTCETLKLQLPNCIEDPVSPIVLRNF	736

QY	721	SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKKNCGNDNI CODDLISITFSFMSLDCLVVG	780
DB	737	SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKKNCGNDNI CODDLISITFSFMSLDCLVVG	796
QY	781	GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV	840
DB	797	GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV	856
QY	841	SGALKSTSCSINHPIFFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
DB	857	SGALKSTSCSINHPIFFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	916
QY	901	QLELPKVAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV	960
DB	917	QLELPKVAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV	976
QY	961	RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELKAPVNVVNCISIAVCQRIQC	1020
DB	977	RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELKAPVNVVNCISIAVCQRIQC	1036
QY	1021	FFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAELLFNDVSFTLLPQCGAFVRSQTET	1080
DB	1037	FFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAELLFNDVSFTLLPQCGAFVRSQTET	1096
QY	1081	KVEPPEVNPPLPLIVGSSVGGLLLALLITALYKLGFFKRYQKDMMSGGPPGAEPPQ	1137
DB	1097	KVEPPEVNPPLPLIVGSSVGGLLLALLITALYKLGFFKRYQKDMMSGGPPGAEPPQ	1153

RESULT 5

AAO14428

ID AAO14428 standard; protein; 1153 AA.

AC AAO14428;

DT 03-MAY-2002 (first entry)

XX Integrin Mac-1 alpha subunit.

DE Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.

OS Unidentified.

XX WO200204521-A2.

PN 17-JAN-2002.

PD 09-JUL-2001; 2001WO-US21805.

XX 07-JUL-2000; 2000US-216600P.

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (BLOO-) CENT BLOOD RES.

PI Springer T;

DR WPI; 2002-148167/19.

XX New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders

PS Example 1; Fig 1F; 90pp; English.

XX The invention comprises structurally biased variant integrin inserted (I)
CC domain proteins, wherein the alterations to the protein occur in at least

two noncontinuous regions. Specifically the the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating: ischaemia/ reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin.

Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 23; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLPFLGSLNRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLPFLGSLNRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMVSTVMEQLKSKTFLSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMVSTVMEQLKSKTFLSLMOYSEEP 196

QY 181 RHFTFKFQNNPRLSKPTTOLLGRTHATGLRVRELFNITNGARKNAFKILFL 240
DB 197 RHFTFKFQNNPRLSKPTTOLLGRTHATGLRVRELFNITNGARKNAFKILVVI 256

QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKRSQELNTVASKPRDRHVFQIN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKRSQELNTVASKPRDRHVFQIN 316

QY 301 NPEALKTONQLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NPEALKTONQLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKESTFFINMTVDSDMDAYLGAAAIILNRVOSLVLAGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESTFFINMTVDSDMDAYLGAAAIILNRVOSLVLAGAPRYOHIGLVAMFR 436

QY 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEYQTRGGQVSVCP 480
DB 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEYQTRGGQVSVCP 496

QY 481 PRGQARWQCDVAVLGEQCPMGFRFGAALTVLGDVNGDKLTDAIGAPGEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLGEQCPMGFRFGAALTVLGDVNGDKLTDAIGAPGEDNRGAVYLF 556

QY 541 HCTSGSGISPSHSORIASKLSPLRYFGOSLSGGDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HCTSGSGISPSHSORIASKLSPLRYFGOSLSGGDLTMDGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLVRKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVT 660
DB 617 PVLVRKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVT 676

QY 661 YDLALDGRPHSRVFNFTKSTRQTVGLTQTCTELKQLPNCIEDPVSPIVLRNLF 720
DB 677 YDLALDGRPHSRVFNFTKSTRQTVGLTQTCTELKQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGLNRLPVLAEDAQRLETFAPPEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNRLPVLAEDAQRLETFAPPEKNCNDNICODDLSITFSFMSLCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 840

DB 797 GPREFNVTVVRNDGEDSYRTQVTFPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 916
QY 901 QLELPVKYAVYVMTVTSHGVSSTKYNFTASENTSRVMQHVQVSNLQSRSLPISLVLPVPV 960
DB 917 QLELPVKYAVYVMTVTSHGVSSTKYNFTASENTSRVMQHVQVSNLQSRSLPISLVLPVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCORICDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCORICDIP 1036

QY 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDISVFTLLPGOGAFVRSOTET 1080
DB 1037 PFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDISVFTLLPGOGAFVRSOTET 1096

QY 1081 KVBPPFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMHSEGGPPGAEPO 1137
DB 1097 KVBPPFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMHSEGGPPGAEPO 1153

RESULT 6
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AC AAR04136;
XX DT 25-MAR-2003 (updated)
XX DT 07-SEP-1990 (first entry)
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
XX synthetic.

Key Location/Qualifiers
FT Modified-site 86..88 /*label= putative N-glycosylation site
FT Modified-site 240..242 /*label= putative N-glycosylation site
FT Modified-site 391..393 /*label= putative N-glycosylation site
FT Modified-site 469..471 /*label= putative N-glycosylation site
FT Modified-site 693..695 /*label= putative N-glycosylation site
FT Modified-site 697..699 /*label= putative N-glycosylation site
FT Modified-site 735..737 /*label= putative N-glycosylation site
FT Modified-site 802..804 /*label= putative N-glycosylation site
FT Modified-site 881..883 /*label= putative N-glycosylation site
FT Modified-site 901..903 /*label= putative N-glycosylation site
FT Modified-site 912..914 /*label= putative N-glycosylation site
FT Modified-site 941..943 /*label= putative N-glycosylation site
FT Modified-site 947..949 /*label= putative N-glycosylation site
FT Modified-site 979..981 /*label= putative N-glycosylation site
FT Modified-site 994..996 /*label= putative N-glycosylation site
FT Modified-site 1022..1024 /*label= putative N-glycosylation site
FT Modified-site /*label= putative N-glycosylation site

FT Modified-site 1045..1047
 FT /*label= putative N-glycosylation site
 FT Modified-site 1051..1053
 FT /*label= putative N-glycosylation site
 FT Modified-site 1076..1078
 FT /*label= putative N-glycosylation site
 FT region 1..16
 FT /*label= signal peptide
 FT region 1106..1134
 FT /*label= putative transmembrane region
 PN EP364690-A.
 XX
 PD 25-APR-1990.
 XX
 PF 17-AUG-1989; 89EP-0115159.
 XX
 PR 23-AUG-1988; 88US-0235353.
 PR 09-MAR-1989; 89US-0321239.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Springer TA, Corbi A;
 XX WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 PT New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 XX
 PS Disclosure; Page 7; ?pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1153 AA;

Query Match 99.3%; Score 5842; DB 11; Length 1153;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVWVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVWVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSIGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLFGSNLRQOPQK 120
 DB 77 RLQVPVEAVNMSIGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLFGSNLRQOPQK 136
 QY 121 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKWVSTVMEQLKSKTLFSLMOYSEEF 180
 DB 137 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKWVSTVMEQLKSKTLFSLMOYSEEF 196
 QY 181 RIHFTKPEQNNPNRSLKIPITQLGRTHATGLRKVVRELFNITNGARKNAKIFLVL 240
 DB 197 RIHFTKPEQNNPNRSLKIPITQLGRTHATGLRKVVRELFNITNGARKNAKIFLVL 256
 QY 241 TDGEKFGDPLGVEDVPELDREGVIRYVIGVDGAPRSEKSRQELNTVASKPRDHVQFN 300
 DB 257 TDGEKFGDPLGVEDVPELDREGVIRYVIGVDGAPRSEKSRQELNTVASKPRDHVQFN 316
 QY 301 NFEALKTIONQREKIFAIEGTQSGSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 360
 DB 317 NFEALKTIONQREKIFAIEGTQSGSSFEHMSQEGFSAATISNGPLLSVGSVDWAG 376
 QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVLGAPRYQHIGLVAMPR 420
 DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLICAPHYYQTRGGQVSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLICAPHYYQTRGGQVSVCP 496
 QY 481 PRGQARWQCDAVLYGEGQGWGRFGAALTIVLGVNGDKLTDVAIGAPBEDNRGAVYLF 540
 DB 497 PRGQARWQCDAVLYGEGQGWGRFGAALTIVLGVNGDKLTDVAIGAPBEDNRGAVYLF 556
 QY 541 HGTSGSISPSHSORISAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTIVGAGHVLRLRSQ 600
 DB 557 HGTSGSISPSHSORISAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTIVGAGHVLRLRSQ 616
 QY 601 PVLRVKAIMBENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQTQSVVT 660
 DB 617 PVLRVKAIMBENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQTQSVVT 676
 QY 661 YDLALDSGRPHSRVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 677 YDLALDSGRPHSRVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
 DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSPLISLVFLVPV 960
 DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSPLISLVFLVPV 976
 QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIQCDDIP 1020
 DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIQCDDIP 1036
 QY 1021 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFAVRSQTET 1080
 DB 1037 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFAVRSQTET 1096
 QY 1081 KVEPEFVNPNPLIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGAEPP 1137
 DB 1097 KVEPEFVNPNPLIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGAEPP 1153

RESULT 7

AAR07120

ID AAR07120 standard; protein; 1163 AA.

XX

AC AAR07120;

XX

DT 25-MAR-2003 (updated)

DT 05-FEB-1991 (first entry)

XX

DE p150.95 alpha subunit encoded by clone lambdaX47.

XX

KW p150, 95 leucocyte adhesion receptor alpha-subunit;

KW hairy cell leukaemia; rhinovirus.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 61..63

FT /label-glycosylation site

FT Modified-site 89..91

FT /label-glycosylation site

FT Modified-site 385..387

FT /label-glycosylation site

OS Homo sapiens.
 XX W0200029446-A1.
 XX 25-MAY-2000.
 PD 16-NOV-1999; 99WO-US27139.
 XX 16-NOV-1998; 98US-0193043.
 XX 08-JUL-1999; 99US-0350259.
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 XX WPI; 2000-387751/33.
 XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous
 PT system injury sites
 XX Example 5; Fig 1; 270pp; English.
 XX Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit; the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This
 CC sequence was used in an alignment to identify a novel beta2 integrin
 CC alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha d. The
 CC Alpha d gene and protein may be useful in therapy for diseases linked
 CC to aberrant alpha d function (e.g. Type 1 diabetes, atherosclerosis,
 CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
 CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
 CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
 CC antibodies may be used in the inhibition of macrophage infiltration at
 CC the site of a central nervous system injury. The monoclonal antibodies
 CC can also be used to detect and diagnose Crohn's disease.
 XX Sequence 1163 AA;
 SQ
 Query Match 58.4%; Score 3436; DB 21; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 4, 1e-278;
 Matches .684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAQNRGSLVQCDSYTGSCPEI 60
 DB 20 FNLDTEULTAFRVDGAGDSVVQVYVANSWVVGAPQKIAAQNIGGLYQCGYSTGACEPI 79
 QY 61 RLOVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 120
 DB 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 137
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
 DB 138 LPVSQECPRQSDIVFLIDGSGSISSRNFMFVRAVISQFQRPSTQSLMQPSNKF 197
 QY 181 RIHTFKFQNNPRSLKIPITQLLGRTHATGLRKVVRLFNITNGARKNAFKILPLL 240
 DB 198 QTHFTFEFRRTSNPLSLASVHQLQGGTYTATAIQNVVHRLFHASYGARRDAIKILVI 257
 QY 241 TDGEKFGDPLGVEDIVPELDREGVTRYVIGVDADRFSKSRQELNTVASKPRDRHVFOIN 300
 DB 258 TDGKEGSDLYQDVIPNADAGIIRYAIQVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
 QY 301 NFEALKTQNLQREKIFAEGTQTCSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 318 DFDALKDTONQLKEKIFAEGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGVSFTWSG 377
 QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGYAAAAIILNRVQSLVILGAPRYQHIGLVAMFR 420
 DB 378 GAPLYPPNMSPTFINNQENVDNRYSYLGYSYTELAKMGVQSLVILGAPRYQHIGKAVIFI 437
 QY 421 QNTGMWESNANVKGTOIGAYFCASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCP 480
 DB 438 QVSRQWRMKAIEVIGTOIGSYFCASLCSVDVDTGSDTLVLIGAPHYEQTRGGQSVCP 497
 QY 481 PRGQARWOCDAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEENRGAVYLF 540
 DB 498 PRGWR-RWMCDAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEENRGAVYLF 556
 QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVCAGQCHVLLRSG 600
 DB 557 HGVLGPSISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLAVGARGOVLLRTR 616
 QY 601 PVLVRKAIMEFNPREVARNVFECDNQVKGKEAGEVVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLWGVSMQFIPAEIPRSFAFECREQVSEQTLVQSNICLYIDKRSKNLGSRLDQSSVT 676
 QY 661 YDLALDSGRPHRAVFNETKSTRQTVLGLTQTCETLKLQPLNCIEDPVSPIVRLNPF 720
 DB 677 LDALAPGRSLSPRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVSDSVPIILRLNF 736
 QY 721 SLVGTPLSAFGLNLRPVLAEDAORLFTALPPEKNKCGNDNICDDLSITTFSEMSLDCLVVG 780
 DB 737 TLVGKPLLAFLNLRPMLAALAQRYFTASLPFEKNCGADHICQNLGISFSPGLKSLVG 796
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRMLACESASSTEV 840
 DB 797 SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSYRYVAEGKQQLASLHLC--CSAPVG 854
 QY 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
 DB 855 SQGTWSTSCRINHLEFRGAQITFLATFDVSPKAVGLDRLLLTANYSSENNIPRTSKTIF 914
 QY 901 QLELPVKYAVYVMTSHGVSTKVLNFTAS-ENTSRVMOHYOVSNLQSRSLPSLVFLVP 959
 DB 915 QLELPVKYAVYVMTSHGVSTKVLNFTAS-ENTSRVMOHYOVSNLQSRSLPSLVFLVP 974
 QY 960 VRLNQTVINDRPQVTFSENLSTCTHKERLPSHDFLAELRKAPVNVCSIAVCQICDI 1019
 DB 975 VELNQEAVMDVEVSHVHPQNPQLSACSEKIAPPASDFLAHQKPNVLDCSIAGCLRPRCDV 1034
 QY 1020 PFGIOQEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1079
 DB 1035 PSFSVQBELDFTLKGNLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1094
 QY 1080 TKVEPEVPNPPLPLIVGSSVGGLLALITAAALYKLGFFKQYKDMWSE 1128
 DB 1095 TVLEKYKVNPIPLIVGSSVGGLLALITAAALYKLGFFKQYKDMWSE 1143
 RESULT 10
 ABG61470
 ID ABG61470 standard; Protein; 1163 AA.
 XX AC ABG61470;
 XX AC AC
 XX 27-AUG-2002 (first entry)
 XX Human Beta2 integrin alphaCD11c subunit.
 DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
 XX LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 XX Homo sapiens.
 OS

PF 08-APR-2002; 2002WO-US10824.
XX
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
XX WPI; 2003-058520/05.
DR
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
XX
PS Claim 1; Page 225-228; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;
Query Match 58.4%; Score 3434; DB 24; Length 1163;
Best Local Similarity 60.5%; Pred. No. 6.1e-278;
Matches 683; Conservative 140; Mismatches 300; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVVGAPQEIIVAAHQRSGLQCYDSTGCEPI 60
DB 20 FNLDTEELTAFRVDSAGSDSVVQVANSWVVGAPQKITAANQTGGLYQCYSTGACEPI 79
QY 61 RLQVPVEAVNMLGLSLAATTPSPQLLACGPTVHTQCSNTYVVKGLCFGLSGNLRQOQPK 120
DB 80 GLQVPPEAVNMLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQESDIAFLIDGSGSIIPHDFFRMKEWVSTMVQLEKSKTLFSLMOYSEEF 180
DB 138 LPVSRQECBPQSDIVFLIDGSGSISSRNFAVMNFVRAVISQFQRPSTQFSLMOFSNKF 197
QY 181 RIHFTFKPQNPNRSLIKPTQLIGRTHATGLRKVVRELFINITNGARKNAFKILFL 240

RESULT 12

AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

198 QTHLTFFEFRTSNPLSLASVHQLQGGFTYTATAIQNVVHRLPHASVGRARDATKILIVI 257
241 TDGEKGDPLGYEDVILPELDREGVIRVYIGVGFADFRSEKSRQELNLTVASKPPRDHVFQIN 300
258 TDGKKEGDTLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKELNDIASQSEHIFKVE 317
301 NFPAKTIQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLPLSTVGSVDWAG 360
318 DFDALDKDIQTLREKIFPIEGTETSSSPSELEMAQEGFSAVTPDGPVLGAVGSEPTWSG 377
361 GVFLYTSKESKSTINMTRVDSMDNDAYLGAAAILLRNVRVOSLVLCAPRVOHIGLVAMER 420
378 GAFLYPPNMSPTFINMSQENQVDRSDYLGYSTELALWKGVOSLVLCAPRYQHTGKAVIFT 437
421 QNTGMWESNANVAGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGQVSCPL 480
438 QVSRQWRMKAETVGTQIGSYFGPSLCSVDVDSGSTDVLIGPHYYEQTGRGQVSCPL 497
481 PRGORARWQDAVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAYILF 540
498 PRGWR-RWMCDAVLYGEGHPWGRFGAALTVLGVNGDKLTDVVIGAPGEEENRGAYILF 556
541 HGTSGSISPSHSORITAGSKLSRLOYFGOSLGGQDLTMDGLVDLTVGAOGHVLRLSQ 600
557 HGVLPSPISPSHSORITAGSKLSRLOYFGQALSGQDLTQDGLVDLAVGARGQVLLLRTR 616
601 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLRVQKSTRDRREGIQSVWT 660
617 PVLWGVGSMOFIPAEIPRSFAFCEQVVSQTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
661 YDLALDSGRPHSRVAVNETKNSRTRQTVLGLTQCTETLKLQLPNCITDVPVILRLNF 720
677 LDIALDPGLSPRATFOETKNRSLSRVVLGLKAHCENFLLPLSPCEDSVTPITLRLNF 736
721 SLVCTPLSAGNLRPVLAEADAORLFTALFPEKNCNGNDNICODDLSITFESFMSLDCLVVG 780
737 TLVGKPLLAFLRNLRLPMLAADAQRYFTASLFFERNKGADHICQDDMLGIFSFPGLKSLVVG 796
781 GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLVSRKVTSTLQNRQSRSWRLACESASTEV 840
797 SNLELNAEVMWMDGSDSYGTTITFSPHAGLSYRYVAGQKQGLRSLHLTCDSPVG-- 854
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTPEF 900
855 SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTSKTTF 914
901 QLELPVKYAVVMVTSYHGVSTKYLNFPTAS-ENTSRVMQHOYOVSNLQORSPLISLVLPV 959
915 QLELPVKYAVVYTVSSSHQFTKYLNFSESEKESHVAMHRYQVNNVNGORDLPVSINFVWP 974
960 VRLNQTVIWDPRQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDI 1019
975 VELNQEAVMWMDVEVSLPQNPSLRCSSEKSIAGPASDFLAHQKNPVLDCSIAGCLRFRCDV 1034
1020 PFFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGQGAFFVRSOTE 1079
1035 PSFVSQBELDFTLKGNSLFGWVRQILOKKVSWSVAEITFTDTSVYSQLPGQEAFFRAQTT 1094
1080 TKVPEFEPVNPPLPIVGVSSVGGILLALLITAAALYKLGFFKQYKDMNSE 1128
1095 TVLEKYKVHNPTPLIVGSSIGGLLLALLITAVLYKVGFFKQYKEMMEE 1143

KW	Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.
XX	Homo sapiens.
XX	Key
FT	Location/Qualifiers
FT	17..1108
FT	Domain
FT	150..352
FT	/note= "extracellular domain"
FT	Region
FT	/note= "this region is homologous to the insert common to CH1a,b,c and may be a site for interaction with ICAM family proteins"
FT	465..474
FT	Binding-site
FT	/note= "putative cation binding site"
FT	518..527
FT	Binding-site
FT	/note= "putative cation binding site"
FT	592..600
FT	Binding-site
FT	/note= "putative cation binding site"
FT	1109..1128
FT	Region
FT	/note= "transmembrane region"
FT	1129..1161
FT	Domain
FT	/note= "cytoplasmic domain"
XX	WO9517412-A1.
XX	PN
XX	29-JUN-1995.
PD	21-DEC-1994; 94WO-US14832.
XX	XX
XX	05-AUG-1994; 94US-0286889.
PR	23-DEC-1993; 93US-0173497.
PR	XX
XX	(ICOS-) ICOS CORP.
PA	XX
XX	Gallatin WM, Van Der Vieren M;
PI	WPI; 1995-240603/31.
DR	N-PSDB; AAQ91712.
DR	XX
XX	Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential antiinflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.
PT	Claim 7; Page 82-87; 172pp; English.
PS	XX
XX	A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha-TM1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.
CC	CC
CC	The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells.
CC	XX
XX	Sequence 1161 AA;
XX	Query Match 57.8%; Score 3401; DB 16; Length 1161;
XX	Best Local Similarity 59.3%; Pred. No. 3.5e-275;
XX	Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
Qy	1 FNLDENAMTFOENARGQSGVWOLQGSRVVVGAPQETVAANORGSLYQCDYSTGSCPEI 60
Db	17 FNLDVEPTTFQEDAGGQSGVWVGGSRVVGAPLEVAANQTRGLDCAAAATGMCPPI 76
Qy	61 RLQVPEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGLCFGLFSNLRQOPQK 120
Db	77 PLHIRPEAVNMSLGLTAAASNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIIQT 135
Qy	121 FPEALRGCPQSDSDIAFLIDSGSIIPHDPRMKWSTVNEQKSKTFLPSLMQYSEEP 180
Db	136 VPDATPECPHOEMDIVELIDSGSIDQDNFQMKGFQVAVMGQFEGDTTLFALMQYSNLL 195
Qy	181 RIHTEFEFONPNRSLIKPITOLLGRTHTATGLRKVRRELFNITNGARKNAFKILFL 240
Db	196 KIHFTFTQRTSPSQSLVDPIVQLKGLTFTATGILTVVTLQFHHKNGARKSAKKILIVI 255

Qy	241 TDGEKFDGPIGYEDVPELDEGVIRVIVGVGAQPRSEKSRQELNVTASKPRDRHVFOIN 300
Db	256 TDGOKYKPLEYSDVIPAQEKAGIIRYAIGVHAFOGTPARQELNVTASSAPPQDHVKVD 315
Qy	301 NFEALKTIONQREKIPAIETGOTGSSSPSEHMSQEGFSAAITNSGPLLSLVGVSQWAG 360
Db	316 NFAALGSIQOLQEKIYAVEGTQSRASSQFHEMSQEGFTALTMDGLFCAGVSPSWG 375
Qy	361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGAIAAILRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db	376 GAFLYPNMSPFTINMSQENVMDRDSYLGYSYTELALWKGQNLVGLGAPRYQHTGKAVFT 435
Qy	421 QNTGMWESNANVKGTOICAYFGASLCSVDVDNSGSTDLLVIGAPHYEQTRGGQVSCPL 480
Db	436 QVSRQWRKKAETGTOIGSYFGASLCSVDVDNSGSTDLLVIGAPHYEQTRGGQVSCPL 495
Qy	481 PRQARWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db	496 PRQORVQWQCDVLRGEQHPWGRFGAALTVDGVNEDKLIDVAIGAPGQENRGAYLYF 555
Qy	541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAGQGHVLLRSQ 600
Db	556 HGASEGSIPSHSQRIASSQLSPRLQYFGQALSGGQDLTODGLMDLAVGARGQVLLRSL 615
Qy	601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVVRVCLHVQKSTRDRREGIOQSVT 660
Db	616 PVLKVGVMRFPSPVEKAVYRWEKPSALEAGDATVCLTIQKSLDQL--GDIQSSVR 673
Qy	661 YDLALDSGRPHSRAVENETNSTRTQVGLTQTCETLKLQPNCLIEDVSPVILNLF 720
Db	674 FDLALDPRLTSRAIFNETKNPTLTRKKTGLGHCHCETLKLUPDCVEDVVSPIILNLF 733
Qy	721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPFPEKNCNDNICQDDLSITFSMSLDCLVVG 780
Db	734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGVTLSFGSLQTLTVG 793
Qy	781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db	794 SSLELNIVTVMNAGDSYGTVSVLYYPAGLSHRRVSGAQKQPHQSLRLACETV-PTED 852
Qy	841 SGALKSTSCSINHPFIPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db	853 EG-LRRSRCSNVNHPFIFHEGSGNGTFIVTFDVSYKATLDGRMLMRASSENKASSSKATF 911
Qy	901 QLELPVYAVYVMVVTSHGVSTKYLNFTASENTRVVMQHOVQVSNLQORSIPISLVLVP 959
Db	912 QLELPVYAVYVMISRQESTKYNFATSDKWKKEAHEHYRVNNSLQORDLAISINFVP 971
Qy	960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
Db	972 VLLNGVAVMDVMVMEAPSQL-PCVSRKPPQSHSDFLTQISRSPLMDCSIADCLQPRCDV 1029
Qy	1020 PFGIQEEFNATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGCGAFVRSQTE 1079
Db	1030 PSFSVQBELDFTLKGNLSPGFWRETTQKKVLYVSVVAEITFDTSVYSQLPQGAQPMRAQME 1089
Qy	1080 TKVEPPEVPNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKQKDMSE 1128
Db	1090 MWLEDEEVNAIFIMGSSVGALLLLALITATLYKLGFFKRYKHEMLED 1138
RESULT 13	
AAW23049	
ID	AAW23049 standard; Protein; 1161 AA.
XX	AAW23049;
AC	AAW23049;
XX	XX
DT	24-FEB-1998 (first entry)
XX	Human beta 2 integrin alpha d subunit.
DE	XX
XX	Beta 2 integrin alpha d subunit; human; cell migration;
KW	cell adhesion; phagocytosis; diabetes; atherosclerosis;

KW multiple sclerosis; asthma; psoriasis; lung inflammation;
XX acute respiratory distress syndrome; rheumatoid arthritis.

OS Homo sapiens.

XX Key Location/Qualifiers

XX Domain 17..1108

XX Domain /label= Extracellular_domain

XX Domain 1109..1128

XX Domain /label= Transmembrane_domain

XX Domain /note= "homologous to the human CD11c transmembrane

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

XX Domain /note= "domain common to CD11a, CD11b and CD11c"

XX Domain 1129..1161

XX Domain /label= Cytoplasmic_domain

XX Domain 150..352

XX Domain /note= "region homologous to the I (insertion)

XX Domain 150..352

Db 196 KIHFTTQFRTSPSQSLVDPVQLKGLTFTATGTLTVVTLFHHKNGARKSAKILIVI 255
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDHVFOIN 300
Db 256 TDGQKYKDPLEYSVDIPAQKAGIIRYAIGVGHAFQGTARQELNTISSAPPQDPVFKVD 315
QY 301 NFEALKTIQNLREKI PAIEGTOTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDNAG 360
Db 316 NFAALGSQKQLEKIYAVETGTSRASSSPFHEMSQEGFSTALTMGDLGFLGAVGFSWSG 375
QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAIILNRVOSLVLGAPRYQHIGLVAMFR 420
Db 376 GAFLYPPNMSPTFFINMSQENVDMDRDSYLGYSYELALWKGQVNLVLAGPRYQHTGRAVFT 435
QY 421 QNTGWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLVIGAPHYQTRGGQVSVCPDL 480
Db 436 QVSROMRKAKEVTGTIGSYFGASLCSDVDSDGSTDLLILGAPHYQTRGGQVSVCPDL 495
QY 481 PRGORARWQCDALVYGEQOPMGREGAALTVLVDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 496 PRGORVQWQCDALVYGEQOPMGREGAALTVLVDVNGDKLTDVAIGAPGEENRGAVLYF 555
QY 541 HGTSGSISPSHSQRIAGSKLSPRLQYFQSLGSGQDLTMDGLVDTLVGAQGHVLLRSQ 600
Db 556 HGASESGISPSHSQRIASSQLSPRLQYFQSLGSGQDLTMDGLVDTLVGAQGHVLLRSQ 615
QY 601 PVLRYKAIEMFNPREVARNVFECNDQVWKGKAGEVRVCLVHVKSTRDLRLRGQTSVVT 660
Db 616 PVLKVGAMRFPVEKAVRVCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
QY 661 YDLALDSGRPHSRVFNKSTRTQTOVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 674 FDLALDPLRLTSRAIFNETKNTLTKRLKLGIGHCETLKLQLPNCIEDPVPVILRLNF 733
QY 721 SLVGTPLSAFAGNLRLVLAEDAQRULTALPFPPEKNCNDNICODDLSTIFSMCLVCLVVG 780
Db 734 SLVREPIPSQNLRLVLAEGSODLFTASLPPEKNCQDGLCEGLGVTLFSGLQTLTVG 793
QY 781 GPREFNVTVNDEGDSYRTQVTFPPDLDSYKVKVSTLQNRQSRQSRWLACESASSTEV 840
Db 794 SGLSLNVIYVWVNGEDSYGTVSYLYAGLSHRRVSGAKQPHOSALACETV--PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 853 EG-LKSSRCVNHPIFHEGSGNCTFIVTDPVSYKATLGDRLMLRASSENKASSSKATF 911
QY 901 QLELPVKYAVYVMVTSVSHGVSTKYLNF--TASENTSVMQHYQVSNLQGRSLPISLVFLVP 959
Db 912 QLELPVKYAVYTMISROEESTKYFNFATSDKKMKEAHRVYVNNLSQRDLAISINFWVP 971
QY 960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQICDI 1019
Db 972 VLLNGVAVWVYMEAPSQSL--PCVSEKPPQHSDFLTQISRPMPLDCLQACLPQRCDV 1029
QY 1020 PFFGQIEEFENATLKGNLSPDVIKTSNHLNLLVSTAEILFNDSVFTLLPQCGAFVRSQTE 1079
Db 1030 PFSVQEEELDTLKGNLSPGVWRETLQKVLVSVVAEITFTSVSYQLPQGEAFMQAOME 1089
QY 1080 TKVEPEVPNPPLVIVSGVGLLALITALYKLGFPFKRQYKDMME 1128
Db 1090 MVLEEDVYNAIPIIMSSVGALLLALITATLYKLGFPFKRHYKEMLED 1138
RESULT 14
AAW72825
ID AAW72825 standard; Protein; 1161 AA.
XX AAW72825;
AC AAW72825;
XX 19-JAN-1999 (first entry)
XX Human alpha-d.

Query Match 57.8%; Score 3401; DB 18; Length 1161;
Best Local Similarity 59.3%; Pred. No. 3.5e-275;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFQENARGGVSVOQSGSRVVVCGAPQEIIVANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDVEPTTFQEDAGFGGVSVOQSGSRVVVCGAPQEIIVANORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVAVNMSLGLSAAATSPQLACGFTVHQTCSNTYVKGCLFFLGSNLRQOPQK 120
Db 77 PLHIRPEAVNMSLGLTIAASTNGSRLAACGPTLHRVCGENSYKSGSCLLIGSRW-EIIQT 135
QY 121 FPEALRCPPEDSDIAPLDGSSGIIPHDPRKKEWSTVMEQLKSKTLFSLMQVSEEF 180
Db 136 VPDATPCPHQEMDIVELIDGSGSIDQNDENQMGKFVQAVMGQFEGDTDLFALMQYSNLL 195
QY 181 RIHFTKFEQNNPNRSLIKPITOLLGTRTHATCLRKVKVRELFNITNGARKNAFKILFL 240

XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.
 XX OS Homo sapiens.
 XX PN US5831029-A.
 XX PD 03-NOV-1998.
 XX PF 07-JUN-1995; 95US-0482293.
 XX PR 07-JUN-1995; 95US-0482293.
 XX PR 23-DEC-1993; 93US-0173497.
 XX PR 05-AUG-1994; 94US-0286889.
 XX PR 21-DEC-1994; 94US-0362652.
 XX PA (ICOS-) ICOS CORP.
 XX PI Gallatin WM, Van Der Vieren M;
 XX WPI; 1998-609318/51.
 XX DR N-P8DB; AAV67281.
 XX PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 XX immunohistochemical analysis
 XX PS Example 5; Column 61-66; 106pp; English.
 XX CC The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2); (5) antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX SQ Sequence 1161 AA;

Query Match 57.8%; Score 3401; DB 19; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 3.5e-275;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDENAMTFOENARGQSVVQLQGRVVGAPQETVAANORGLYCDYSTGSCPEI 60
 DB 17 FNLDVEEPTIQQEDAGGFGQSVVQVGGSRVVGAPLEVVAAQTGRLYDCAATGMCOPI 76

QY 61 RLQPVVEANMSLGLSLAATSPOLLACGPTVHTQTSNTYVKGCLFGLSNLRQOPK 120
 DB 77 PLHIREAVNMSLGLTAASTNGSLACGPTLHRVCGENSYSGCLLLGSRW-EITQT 135

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEWSTVMEQLKSKTFLSLMQYSBEF 180
 DB 136 VPDATPECPHQEMDIVFLIDGSGSIDQDNFMQMGFVQAVMGQFEGTDTLFPALMQYSNLL 195

QY 181 RIHETFEKFNPNPRSLIKPITQLLGRTHATGLRKVVRELFTNITNGARKNAKIFLL 240
 DB 196 KIHFTFTQFTSPSQSLVDPIVLQKGLTFTTGLTITVTLQFHHKNGARKSAKKILIVI 255

QY 241 TDGEKFGPLGVEDVPELDEGVIRYVIGVDAPRSEKSKQELNTVASKPRDHVFOIN 300
 DB 256 TDGQYKDPLEYSDVIPQAEKAIIRYAIQVGHAFQGTARQELNTISSAPPQDHVFKVD 315

QY 301 NFELKTIQNLREKIFAETQGTQSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 375

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILRNRRVQSLVGLGAPRYQHIGLVAMER 420
 DB 376 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSSTELALWKGQNLVLGAPRYQHTKAVIFT 435

QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTQGGQVSCPL 480
 DB 436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTQGGQVSCPL 495

QY 481 PRGORARWQCDVLYGEGOPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
 DB 496 PRGORVQWQCDVLRGEQHPWGRFGAALTIVLGDVNDKLDIDVAIGAPGOENRGAYLYF 555

QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGOVLLRLSL 615

QY 601 PVLKVKAIMFNPREFARNVPECNDQVVKGEKEGVRVCLHVQKSTDRRLRBEQIQSVVT 660
 DB 616 PVLKVGVMARFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

QY 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
 DB 674 FDLALDPGLRTSRAIFNETKNPTLTRRKTIGLGIHCETLKLUPDCVEDVVSPIILHLNF 733

QY 721 SLVGTPLSAFONLRPVLAEADAQRLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGQDLFTASLPFEKNCQDGLCEGDLGVTLSFGSLQTLTVG 793

QY 781 GPREFNVTVRNDGDSYRTQVTFPPDLDSVRKVTSTLQNRQSRQSRWLACESASSTEV 840
 DB 794 SSLELNIVTVNAGDSYGTWVSLYYPAGLSHRRVSGAQKQPHOSALRLACETV-PTED 852

QY 841 SGALKSTSCSNHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 DB 853 EG-LRSRCSNVNHPHFHEGSGNGTFTVFDVSYKATLDGRMLRASASNNKASSKATF 911

QY 901 QLELPVYAVYVMTSHGVSTKLYNF-TASENTRVMQHOVQVSNLQORSUPISLVFLVP 959
 DB 912 QLELPVYAVYVMTSROESTKYFNFATSDKKWKEAHEHYRVNLSQRLDLAISINFVP 971

QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCIDI 1019
 DB 972 VLLNGVAVVDMVMEAPSQL--PCVSRKPPQHSDFLTQISRSPLMDCSIADCLQFRCDV 1029

QY 1020 PFGIGQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTE 1079
 DB 1030 PSFSVQBELDFTLKGNSLFGVRETLQKKVLSVSAEITFTSVYSQLPQGEAFPMQME 1089

QY 1080 TKVEPPEVPNPLPLIVGSSVGGILLALITAAALYKLGKFKROYKDMSE 1128
 DB 1090 MVEEDEVYNAIPIMGSSVGALLLALITATLYKLGFGRHYKEMLED 1138

RESULT 15
 AAW65089
 ID AAW65089 standard; Protein; 1161 AA.
 XX AAW65089;
 AC AAW65089;
 XX
 DT 28-SEP-1998 (first entry)
 XX
 DE Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 XX
 KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 XX OS Homo sapiens.
 XX FH Key
 FT Peptide Location/Qualifiers 1..16

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:17:44 ; Search time 12.9635 Seconds
(without alignments)
4501.575 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKROYKDMSEGGPFGABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 271250 seqs, 51324744 residues

Total number of hits satisfying chosen parameters: 271250

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:**

1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5821.5	98.9	1152	5	US-09-592-617C-43
2	3444.5	58.5	1200	1	PCT-US03-28227-4834
3	3401	57.8	1161	7	US-60-517-843-2
4	3336.5	56.7	1108	1	PCT-US03-28227-4833
5	3308	56.2	1145	1	PCT-US03-28227-4832
6	3050.5	51.8	1047	7	US-60-487-610-1574
7	1546.5	26.3	1170	7	US-60-487-610-1880
8	1543.5	26.2	1170	5	US-09-592-617C-42
9	1166	19.8	413	7	US-60-487-610-1788
10	1144	19.4	1179	7	US-60-487-610-2494
11	1140	19.4	1179	1	PCT-US03-18234-2
12	1103.5	18.8	1149	7	US-60-487-610-1507
13	1093.5	18.6	1179	6	US-10-474-794-250
14	1055	17.9	1180	6	US-10-474-794-307
15	1054	17.9	1179	7	US-60-493-369-8
16	1054	17.9	1179	7	US-60-512-690-8
17	1054	17.9	1181	7	US-60-493-369-10
18	1054	17.9	1181	7	US-60-512-690-10
19	1052	17.9	1181	7	US-60-487-610-2692
20	1052	17.9	1181	7	US-60-485-450-1726
21	1052	17.9	1181	7	US-60-493-369-9
22	1052	17.9	1181	7	US-60-512-690-9
23	1051.5	17.9	1167	7	US-60-490-890-559
24	1050	17.8	1177	6	US-10-461-862-146
25	1048	17.8	1179	6	US-10-461-862-148
26	1043	17.7	1147	1	PCT-US03-00252A-42

ALIGNMENTS

RESULT 1

US-09-592-617C-43

; Sequence 43, Application US/09592617C

; GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin

; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN ANTAGONISTS

; FILE REFERENCE: 00786-267002

; CURRENT APPLICATION NUMBER: US/09/592,617C

; CURRENT FILING DATE: 2000-06-13

; PRIOR APPLICATION NUMBER: US 08/380,167

; PRIOR FILING DATE: 1999-01-30

; PRIOR APPLICATION NUMBER: US 08/216,081

; PRIOR FILING DATE: 1994-03-21

; PRIOR APPLICATION NUMBER: US 07/637,830

; PRIOR FILING DATE: 1991-01-04

; PRIOR APPLICATION NUMBER: US 07/539,842

; PRIOR FILING DATE: 1990-06-18

; PRIOR APPLICATION NUMBER: US 07/212,573

; PRIOR FILING DATE: 1988-06-28

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 43

; LENGTH: 1152

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -16 to -1

; US-09-592-617C-43

Query Match 98.9%; Score 5821.5; DB 5; Length 1152;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY	1	FNLDTENAMTFOENARGFGQSVVQLQGSVVVVGAPQEIIVANQRGSLYQCDYSTGSCBPI	60
DB	17	FNLDTENAMTFOENARGFGQSVVQLQGSVVVVGAPQEIIVANQRGSLYQCDYSTGSCBPI	76
QY	61	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTVYVKGICFLFGSNLRQOPK	120
DB	77	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTVYVKGICFLFGSNLRQOPK	136
QY	121	FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMBQLKSKTLFSLMOYSEBF	180
DB	137	FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBQLKSKTLFSLMOYSEBF	196
QY	181	RIHFTFKFQNNPNRSLIKPITQLLGRTHRTATGURKVVRELFNITNGARKNAFKILFUL	240
DB	197	RIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILVVI	256

241 TDGKFGDPLGYEDVPELDREGVIRVIGVGDFAFRSEKSRQELNVTASKPRPDHVFQIN 300
Db
257 TDGKFGDPLGYEDVPELDREGVIRVIGVGDFAFRSEKSRQELNVTASKPRPDHVFQIN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTREGGQVSCPL 480
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTREGGQVSCPL 496
Qy 481 PRGORARWQCDVILYGEQGPWGRFGAALTVLGVDVNGDKLTDAIGAAGEEDNRGAAYLF 540
Db 497 PRG-RARWQCDVILYGEQGPWGRFGAALTVLGVDVNGDKLTDAIGAAGEEDNRGAAYLF 555
Qy 541 HGTSGSGTSPSHSORIAGSKSLPRLOYFGQSLSGQDITMDGLVDLTVGAGHVLLRSQ 600
Db 556 HGTSGSGTSPSHSORIAGSKSLPRLOYFGQSLSGQDITMDGLVDLTVGAGHVLLRSQ 615
Qy 601 PVLRVKATMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
Db 616 PVLRVKATMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSTRRTQVQLGTQTCETLKLQPNCEIDPVSIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSTRRTQVQLGTQTCETLKLQPNCEIDPVSIVLRNF 735
Qy 721 SLVGTPLSAFNGLRPLVLAEDAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVWG 780
Db 736 SLVGTPLSAFNGLRPLVLAEDAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVWG 795
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db 796 GPRESNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 960
Db 916 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 976 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHDFLAELRKAPVNVCSIAVCORIQCIP 1035
Qy 1021 PFGIOEENATLKGNSLSEFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQDET 1080
Db 1036 PFGIOEENATLKGNSLSEFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQDET 1095
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALKYKGFQKQYKQMMSEGGPPGAEPO 1137
Db 1096 KVEPEVNPPLIIVGSSVGGLLLLALITAAALKYKGFQKQYKQMMSEGGPPGAEPO 1152

RESULT 2

PCT-US03-28227-4834

; Sequence 4834, Application PC/TUS0328227

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
; APPLICANT: MARTANOVIC, Mirjana M.; SHEN, Fan.
; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;

; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO 4834
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 965829.PT40P
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (1200)
; OTHER INFORMATION: unknown or other
PCT-US03-28227-4834

Query Match

58.5%; Score 3444.5; DB 1; Length 1200;

Best Local Similarity 59.1%; Pred. No. 2.4e-191;

Matches 689; Conservative 142; Mismatches 292; Indels 43; Gaps 5;

Qy 1 FNLDTENAMTFQENARFGQSVVVOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 20 FNLDTELTAFRVDSDAGSGDSVVQYANSWVVGAPQKITAAQOTGGYQCGYSTGACPEI 79
Qy 61 RLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLGSLNRQOPOK 120
Db 80 GLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHHCGRNMYLTGLCFLLGPT--QLTOR 137
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKEWSTVMEOLKSKSTLFLSMQXSEEF 180
Db 138 LPVSRQECPCQEQDIIVFLIDGSGSISRNFATMMNFVRAVISQFQRPSTQFSLMQFSNKF 197
Qy 181 RIHFTFKFQNNPNRSLIKPITQLGRTHATGLRKVRELNITNGARKNAFKILFL 240
Db 198 QTHFTPEFRSSNPLSLASVHQLQGTFTATAIQNVVHRLPHASVGARRDAKILIVI 257
Qy 241 TDEKFGDPLGYEDVPELDREGVIRVIGVGDFAFRSEKSRQELNVTASKPRPDHVFQIN 300
Db 258 TDGKFGDSDLDYKDVIPMAAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIPKVE 317
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
Db 318 DFDALDKIQNLQKEKIFAIEGTQTTSSSSFELEWAQEGFSAVTFDGPVGLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVSQSLVGLGAPRYOHTKAVIFT 437
Qy 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTREGGQVSCPL 480
Db 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTREGGQVSCPL 497

QY 481 PRQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEOQHMPGRFGAALTVLGDVNGDKLTDVVI GAPGEEDNRGAVYLF 556
QY 541 HGTSGGSI SPSSHORTAGSKLSPRLOYFGOSLSGGODLMDGLVDLTGCAOQHVLRLSQ 600
Db 557 HGVLPSPISPSHSORTAGSOLSRLOYFGAUSGGODLTODGLVDLAVARGOVLLRLTR 616
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQWVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660
Db 617 PVLWVGVSQWQFPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLQSSVT 676
QY 661 YDLALDSGRPHRAVNETKSTRQTQVLGTLTQCTETIKLOLPNCIEDPVPVILRLNF 720
Db 677 LDALDPGRLSPRATFOETKNRSLSRVRVLGLKAHCENFNLLPSCEVDSVPTILRLNF 736
QY 721 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNICODDLSITFSPMSLCLVVG 780
Db 737 TLVGKELLAFNLRLPMLAADAQRYFTASLPFEKNCAGADHICQDNLGISEFPGKSLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFEPLDLSYRKVSTLQNRORSORSMWLACESASSTEV 840
Db 797 SNLELNAEVMVNDGEDSYRTITFSPHAGLSRYVVAEGOKQQLASLHLTCDASAPVG-- 854
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKYTSENNPRTNKTEF 900
Db 855 SQGTWSTSCRINHPIFRGQAQITFLATFDVSPKAVLGDRLLTANVSSNNTPRTSKTF 914
QY 901 QLELPVKYAVYVWVTSHGVS TKYLNFTAS-ENTSRVNHQY----- 940
Db 915 QLELPVKYAVYVWVSSHEQFTKYNLFSESEKESHVAMHRYAGGGDAGDAGVGGWEPE 974
QY 941 -----QVSNLQORSPLSLVLPVRLNQTIVWDRPQVTFSENLSST 982
Db 975 TGEFGFLWGLQCPCPLXQVNNLQORDLPVSINFVWVPELVNQAVMDVEVSHQPNSLR 1034
QY 983 CHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEENATLKNLSFDWYI 1042
Db 1035 CSSEKIAPASDPLAHIQKNPVLDCSIAGCLRFCDVPSFVQGEELDTLKNLSFGWVR 1094
QY 1043 KTSNHLILVSTRAIILFNDSVFTLLPGQAFVRSQTEKVEFEVFNPIPLIVSSVGL 1102
Db 1095 QILQKRVSVVSAEITFDTSVYQLPQGEAFMRAQTTVLEKYKHNPPLIVSSIGGL 1154
QY 1103 LLLALITALYKLGFFKROYKDMWSE 1128
Db 1155 LLLALITALYKVGFFRQYKEMWEE 1180

RESULT 3

US-60-517-843-2
; Sequence 2, Application US/60517843
; GENERAL INFORMATION:
; APPLICANT: Weaver, Lynne
; TITLE OF INVENTION: METHODS OF TREATING CHRONIC PAIN USING COMPOSITIONS THAT
; TITLE OF INVENTION: SPECIFICALLY BIND CD11D (ALPHA-D) INTEGRIN
; FILE REFERENCE: 27866/35702
; CURRENT APPLICATION NUMBER: US/60/517, 843
; CURRENT FILING DATE: 2003-11-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-517-843-2

Query Match 57.8%; Score 3401; DB 7; Length 1161;
Best Local Similarity 59.3%; Pred. No. 7.6e-189;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60

Db 17 FNLDVEEPTIIFOEDAGGFGQSVVQPGGRLVVGAPLEVAANQTRLVYDCAATGMCQPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLCFGLFSNLRQQPOK 120
Db 77 PLHIRPEAVNMSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIIQT 135
QY 121 FPEALRGCPQEDSDIAFLIDSGSIIPIHDFRMKEWSTVMEQLKKSKTILFSLMOYSREEF 180
Db 136 VPDATPECPHOEMDITVFLIDSGSIDQDNFMQMGFVQAVMGQFEGTDTLTFALMOYSNLL 195
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVURELFNITNGARKNAFKILFL 240
Db 196 KIHFTTQFRTPSQQSLVDPIVQLGDTFTATGLTIVVTQLFHKGARKSAKILIVI 255
QY 241 TDGEFGPLGVEDYIPELDREGVIRYVIGVGDAPFRSEKSRQELMTVASKPPDRHVFOIN 300
Db 256 TDGQYKDPLEYSVDVIPOAEKAGIIRYAIUGVGHAFQGPARTARQELNTISSAPPQDHVFKVD 315
QY 301 NPEALKTIQONLREKI FAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVSGYDWAG 360
Db 316 NPAALGSIQKQLEKIXAVEGTQSRASSSFHEMSQEGFSTALTMGDLFLGAVGSPFWSG 375
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAVLGYAAAIIILNRVQSLVLAGRYOHI GLVAMPR 420
Db 376 GAFLYPPNNMSPPTFINMSQENVDMSYLGYSTELALWKGQVNLVLGAPRYQHTGRAVPT 435
QY 421 QNTGWESNANVKGQIGAYFCASLCSDVDNSGSTDLVLICAPHYEQTRCGQVSVCP 480
Db 436 QVSRQWRKAEVGTQIGSYFCASLCSDVDSDGSTDLILICAPHYEQTRCGQVSVCP 495
QY 481 PRQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 496 PRQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGGSI SPSSHORTAGSKLSPRLOYFGOSLSGGODLMDGLVDLTGCAOQHVLRLSQ 600
Db 556 HGTSGGSI SPSSHORTAGSKLSPRLOYFGOSLSGGODLMDGLVDLTGCAOQHVLRLSQ 615
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQWVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660
Db 616 PVLKVGAMRFPSPVEKAVYRCWEKPSALEAGDATVCLTIQKSSLDQ--GDIOSSVR 673
QY 661 YDLALDSGRPHRAVNETKSTRQTQVLGTLTQCTETIKLOLPNCIEDPVPVILRLNF 720
Db 674 FDLALDPGRLSRAIFNETKNPTLRRKTLGLGHCETIKLLPDCVEDVWSPILHLNF 733
QY 721 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNICODDLSITFSPMSLCLVVG 780
Db 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGLGVTLSFGLQTLTVG 793
QY 781 GPREFNVTVVRNDGEDSYRTQVTFEPLDLSYRKVSTLQNRORSORSMWLACESASSTEV 840
Db 794 SSLELNIVTVWVNDGEDSYRTVSYIYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKYTSENNPRTNKTEF 900
Db 853 EG-LRSSRCSVNHPIFHEGSGNCTFTVTPDVSKATLGDRLMLRASASSENNKASSKATF 911
QY 901 QLELPVKYAVYVWVTSHGVS TKYLNFTAS-ENTSRVNHQYQVSNLQORSPLIVPLVP 959
Db 912 QLELPVKYAVYVWVTSHGVS TKYLNFTAS-ENTSRVNHQYQVSNLQORSPLIVPLVP 971
QY 960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
Db 972 VLLNGVAVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSADCLQRCVDV 1029
QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLILVSTRAIILFNDSVFTLLPFGQAFVRSQTE 1079
Db 1030 PSFVSQEEELDTLKNLSFGWVRETLQKVLVSVVAEITFDTSVYQLPQGEAFMRAQME 1089
QY 1080 TKVEPEVFNPIPLIVSSVGGGLLLALITALYKLGFFKROYKDMWSE 1128

Db 1090 MWLEDEBVYNAIPIMGSSVYCALLLITATLYKLGLFFKRRHYKEMLED 1138

RESULT 4

PCT-US03-28227-4833
 ; Sequence 4833, Application PC/TUS0328227
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
 ; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
 ; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
 ; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
 ; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
 ; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
 ; APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;
 ; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
 ; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
 ; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
 ; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
 ; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
 ; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;
 ; APPLICANT: WU, Mingham C.; STUIVE, Laura L.;
 ; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
 ; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
 ; APPLICANT: VITT, Ureula A.; KIRTON, Edward;
 ; APPLICANT: XU, Yuming; KWONG, Mary;
 ; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
 ; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
 ; APPLICANT: GITZEN, Darryl; PATURY, Srikanth;
 ; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
 ; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: PN-0100 PCT
 ; CURRENT APPLICATION NUMBER: PCT/US03/28227
 ; CURRENT FILING DATE: 2003-09-12
 ; PRIOR APPLICATION NUMBER: US 60/410,260
 ; PRIOR FILING DATE: 2002-09-12
 ; PRIOR APPLICATION NUMBER: US 60/410,259
 ; PRIOR FILING DATE: 2002-09-12
 ; NUMBER OF SEQ ID NOS: 5444
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 4833
 ; LENGTH: 1108
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 965829.PT35p
 PCT-US03-28227-4833

Query Match 56.7%; Score 3336.5; DB 1; Length 1108;
 Best Local Similarity 59.3%; Pred. No. 3.9e-185;
 Matches 670; Conservative 128; Mismatches 270; Indels 61; Gaps 5;
 Qy 1 FNLDTENAMTQENARFGSGVVOQSGSRVVVGPBEIVANORGSLYQCDYSTGSCPEI 60
 Db 20 FNLDTBELTAPRVDSAGFGSVQVYANSVWVVGAPQKITAANOTGGLYQCCYSTGACEPI 79
 Qy 61 RLQVPVAVNMSGLSLAATTSPPOLLACGTVHTCSENTYVVGICFLGSGNLRQPOK 120
 Db 80 GLQVPPVAVNMSGLSLAATTSPPOLLACGTVHTCSENTYVVGICFLGSGNLRQPOK 137
 Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
 Db 138 LPVSRQECPREQDIIVLIDGSGSISSENFATMNFVRAVISQFQSTQSLMPSNKF 197
 Qy 181 RIHTTFKEFQNNPRSLKIPITQLGRTHATGLRKVKVRELFINITNGARKNAKPIELLL 240
 Db 198 QTHFTFEFRSSNPLSLASVHQLQGTYYTATAIQNVVHRLPHASVARRDAKILIVI 257
 Qy 241 TDGKFGDPLGYEDVIELDREGVIRYVIGVGDFAERSEKSKQELNTVASKPPRDHVFQIN 300
 Db 258 TDGKKEGSDLYDKDVIWMAADAGIIRYAIGVGLAFQNRNWSKELNDIASKPSQSHIFKVE 317
 Qy 301 NFEALKTQNLQREKIFAIEGTQTTGSSSSFEHEMSQEGFSAAITSNGLPILISTVGSYDWAG 360

Db 318 DFDALKDIOQLKEKIFAIEGTETTSSSSFELEMAQEGPSAVFTPDGVLGAVGSPWSG 377
 Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILRNRRVQSLVILGAPRYQHILGVAMFR 420
 Db 378 GAFLYPPNMSPTFINNSQENVDMSYLGSTELALWKGVSQSLVILGAPRYQHTGKAVFT 437
 Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYYEOTRGQGVSCPL 480
 Db 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSNGSTDLVILGAPHYYEOTRGQGVSCPL 497
 Qy 481 PRQRARWOCDAVLGYEQGPWGRFOAALTVLGDVNGDKLTDAICAPGEEDNRGAVYLF 540
 Db 498 PRGWR-RWMCDAVLGYEQGPWGRFOAALTVLGDVNGDKLTDAICAPGEEDNRGAVYLF 556
 Qy 541 HGTSGSGISPSHSQRTAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTUAGQGHVLLRSQ 600
 Db 557 HGVLPSPISPSHSQRTAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTUAGQGHVLLRSQ 614
 Qy 601 PVLVRKAIMFNPREVARNVFCNDQVWKGKEAGEVRVCLHVQKSTRDLRREQIOISVVT 660
 Db 615 -----DLQSSVT 621
 Qy 661 YDLALDSGRPHSRVNETKNSTRQTQVGLTQTCTETKLQPLNCIEDPVPSPVIRLNF 720
 Db 622 LDALDPGRLSPRATFOETKNRSLSRVRLGLKHCENFNLLPSCVEDSVTPIRLNF 681
 Qy 721 SLVGTPLSAFNLRLPVLAEQAORLTALPFENKCGNDNICODDLSITFSPLSLDCLVWG 780
 Db 682 TLVGKPLLAFLRNLRLPMLAADAORYFTASLPFENKCGADHICODNLSIFSPLGLKLLVG 741
 Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPLDLSYRKVSTLQNRORSQSWRLACESASSTEV 840
 Db 742 SNLELNAEVMVNDGDSYRTQVTFPPLDLSYRKVSTLQNRORSQSWRLACESASSTEV 799
 Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTET 900
 Db 800 SGTWSTSCRINHLPFGGAQITFLATFDVSPKAVLGDRLLLTANVSSENMTPTSKTTF 859
 Qy 901 QLELPVKYAVYVTVSHGVSTKYLNFTAS-ENTSRVMQHOYQVNSLQORSLPSISLVLP 959
 Db 860 QLELPVKYAVYVTVSHGVSTKYLNFTAS-ENTSRVMQHOYQVNSLQORSLPSISLVLP 919
 Qy 960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQICQI 1019
 Db 920 VELNQBAVMDVEVSHQPNSLRCSSEKIAPPASDFLAHQKPNVLDCSIAGCLRFRCDV 979
 Qy 1020 PFFGIQEFNATLKNLSFDWYIKTSHNHLIIVSTAEIILFNDSVFTLLPQGGAFVRSQTE 1079
 Db 980 PPSVQBELDFTLKNLSFGWVRQILQKKVSVSVVAEITPDTSVYSQLPQGEAFMRAQTT 1039
 Qy 1080 TKVPEPEVNPPLPIVGVSSVGGLLALITAAALYKLGFFKRYKQVDMWSE 1128
 Db 1040 TVLEKYKVNPTPLIVGVSSVGGLLALITAAALYKLGFFKRYKQVDMWSE 1088

RESULT 5

PCT-US03-28227-4832
 ; Sequence 4832, Application PC/TUS0328227
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;
 ; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;
 ; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan;
 ; APPLICANT: HARTSHORNE, Toinette A.; SUCHOROLSKI, Martin;
 ; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;
 ; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;
 ; APPLICANT: DELEGANE, Angelo M.; PANESAR, Iqbal S.;
 ; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;
 ; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;
 ; APPLICANT: PANZER, Scott R.; WANG, Xinhao;
 ; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;
 ; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;
 ; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;

```
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Barry; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO: 4832
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 965829.PT34p
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (1145)
; OTHER INFORMATION: unknown or other
; PCT-US03-28227-4832

Query Match      56.2%; Score 3308; DB 1; Length 1145;
Best Local Similarity 57.5%; Pred. No. 1.8e-183;
Matches 670; Conservative 128; Mismatches 270; Indels 98; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVANQROSLVOCYSTGSCPEI 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
20 FNLDTTELTAFRVDSAGFGDSVVQVANSVVVVGAPQKITAANQTLGGLYCGYSTGACEPI 79
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQOQPK 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
80 GLQVPPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCLLGT.-QLTOR 137
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKKSKTLFSLMOYSBEF 180
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
138 LPVSQECPRQDQIVFLIDGSGSISSRNFAFMFVRAVISQFORPSTQSLMGFSNKF 197
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
181 RIHFTKFPQNNPRSLKIPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
198 QTHFTFEEPRSSNPLSLASVHQLQGFTYTATQNVVHRLFHASYGARRDAAKILIVI 257
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGUGDAFRSEKSRQELNTVASKPPRDHVQIN 300
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
258 TDGKEGSLDYKDVIPADAAIIRYAIYGVGLAFQNRNSWKLNDIAKPSQEHFKVE 317
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
301 NFEALKTIONQLREKIFALEGTQGTSSSFEHMSQEGFSAIITNSGPLLSTVGSVDWAG 360
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
318 DFDAKLDIQNKEKIFALEGTETSSSFEHMSQEGFSAVFTPDGVLGAVGFTWSG 377
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
361 GVFLYTSKEKSTFINMTRVDSMDNAYLGAAIIRNRVQSLVGLGAPRYQHIGLVAMFR 420
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALMKGVQSLVGLGAPRYQHTGRAVFT 437
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
421 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCLP 480
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
438 QVSRQWRMKAETVGTQIGSYFGASLCSDVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCLP 497
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
481 PRGQARWOCDAVLVGEQGPWGRFGAALTVLGDVNGDKLTDVAICAPGEENRGAVLYF 540
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
498 PRGWR-RWMCDAVLVGEQGPWGRFGAALTVLGDVNGDKLTDVTCAPGEENRGAVLYF 556
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
541 HCTSGSGISPSHRSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 600
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Barry; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO: 4832
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 965829.PT34p
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (1145)
; OTHER INFORMATION: unknown or other
; PCT-US03-28227-4832

; DB 557 HGVGPSISPSHRSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 614
; QY 601 PVLAVKAIMEFNPREVARNVPECNDQVVVKGKEAGEVRVCLHVQKSTRDLRLEAGIQSVVT 660
; DB 615 -----DLQSSVT 621
; QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
; DB 622 LDALDLPGLSRPRATFOETKNSRSLRVRLGLKAKCENFNLLPSCVEDSVTPITRLNF 681
; QY 721 SLVCTPLSAFCNLRPVLAEDAORLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
; DB 682 TLVGKPLLAFLNLRPLMLAADAQRYTASLPPPEKNCAGADHICODNLGISFSFPLGKSLDVG 741
; QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
; DB 742 SNLELNAEVMVWNGEDSYGTTITFHPAGLSYRYVAEGQKQGLRSLHLTCDAPVG-- 799
; QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
; DB 800 SGTWSTSCRINHILIFRGAQITFLATPDVSPKAVLGDRLLLTANVSSENTPRTSKTTF 859
; QY 901 QLELPVKYAVYVTVSHCVSTKYLNTAS-ENTSRVMOHOY----- 940
; DB 860 QLELPVKYAVYVTVSHCVSTKYLNTAS-ENTSRVMOHOY----- 940
; QY 941 -----QVSNLGRSLPISLVLPVRLNQTIVLWDRPQVTFESNLST 982
; DB 920 TGEFGFLWRGLCPQCPPLXQVNNLQORDLPVINFVPELQEAVMVDEVSHQPNPBLR 979
; QY 983 CHTKERLPSHSDFLAELKAPVWNCISIAVCQRIQCDIPFGIQQEENFATLKNLSFDWYI 1042
; DB 980 CSSEKIAPPASDFLAHIQKNPVLDCSIAQCLFRCDVPSFVQEBELDTLKNLSFGWVR 1039
; QY 1043 KTSNNHLLIVSTAELFNDSVFTLLPGQAFVRSQETKTEPFEFVNPPLPLIVGSSVGL 1102
; DB 1040 QILQKQSVWSVAGITFDTSYQLPGQEAFMRAQTTTLEKYKVHNTPPLIVGSSVGL 1099
; QY 1103 LLLALITAAALYKGLGFFKROYKDMSE 1128
; DB 1100 LLLALITAVLYKVGFFKROYKEMEE 1125

RESULT 6
US-60-487-610-1574
; Sequence 1574, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1574
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1574

Query Match      51.8%; Score 3050.5; DB 7; Length 1047;
Best Local Similarity 54.7%; Pred. No. 1.2e-168;
Matches 610; Conservative 146; Mismatches 256; Indels 103; Gaps 5;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVANQROSLVOCYSTGSCPEI 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
18 FNLQDVEEPTIFQEDAGGFGQSVVQFGGRLVVGAPLEVAANQTLGRLLDCAATGMCQPI 77
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQOQPK 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuming; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Barry; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO: 4832
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 965829.PT34p
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (1145)
; OTHER INFORMATION: unknown or other
; PCT-US03-28227-4832
```


Db 719 QDLISPINVSNFSLWEEBGTDRQRAQKIDPILRPSLHSETWEI-----PFERKNCGE 773
QY 758 DNIQODDLSITFSFMSLDCLVGGPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVS 817
Db 774 DKKEANLRVSFSPARSRALRLTAFASLSVELSLNLEEDAYVWQLDLHFPGLSPRKVE 833
QY 818 TLQORSORSWLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTVNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCCELPESRLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887
QY 876 LGNKLKLLKANVTSENN---MPTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNTFASEN 931
Db 888 WGSVELHANVTCNNEDSLLLEDNSATTI---IPILYPINILIOQOEDSTLYVSFTPKGP 944
QY 932 TSRVMQHOYQV---SNLQORSPL-ISLVFLVPLRLNQTIVWRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKMYQVRIQPSIHNDIPTLEAVGVGPQPPSEGPITHQWSVQMEPPV--PCHVED 1002
QY 987 -ERLPKSHSD--FLAELRKAPVNCISAVCQRIQCDIPFGIOEBFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAAEPCLPGALFRCPVW-----FRQELVQVIGTLELVGEIE 1044
QY 1044 TSHNHLIVSTAELFNDSVFTLLPGOGAFVRSOTETKVEPEVNPPLPLIYVSSVGGILL 1103
Db 1045 AS-SWFSLCSSLSISFNSSKHFLYGSNASL-AQVMKVDDVYVYKQMLYLYVLSGIGILL 1102
QY 1104 LLALITAALYKLGFFKRYQKDMMSSEG-GPPGAEP 1136
Db 1103 LLLLIFIVLYKGVFFKRNLEKMEAGRGVNGIP 1136

RESULT 8

US-09-592-617C-42
; Sequence 42, Application US/09592617C
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Armin
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN ANTAGONISTS
; FILE REFERENCE: 00786-267002
; CURRENT APPLICATION NUMBER: US/09/592,617C
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 08/380,167
; PRIOR FILING DATE: 1999-01-30
; PRIOR APPLICATION NUMBER: US 08/216,081
; PRIOR FILING DATE: 1994-03-21
; PRIOR APPLICATION NUMBER: US 07/637,830
; PRIOR FILING DATE: 1991-01-04
; PRIOR APPLICATION NUMBER: US 07/539,842
; PRIOR FILING DATE: 1990-06-18
; PRIOR APPLICATION NUMBER: US 07/212,573
; PRIOR FILING DATE: 1988-06-28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 42
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-592-617C-42

Query Match 26.2%; Score 1543.5; DB 5; Length 1170;
Best Local Similarity 34.3%; Pred. No. 1.4e-81;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;
QY 1 FNLDTENAMTFQ--ENARGFGOSVVOQGSRVVVGAPQEIIVANQORGSLYQCDYSTGSC 58
Db 26 YNLDVRGARSFPFRAGRHFGRVLOV--GNGVIVGAPGE---GNSTGSLYQCSQGTGHL 81
QY 59 PIRLQVPVAVNMSLGLSLAATSPQLIACGPTVHQTCSENTYVKGLCFRGSNLR--- 115
Db 82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFRQNLQGM 140
QY 116 -QOQKFPFALRGCGQEDSDIAFLIDGSGSIIPHDFRNMKEWSTVMEOLKSKTLFSLM 174
Db 141 LQGRPGFQRCIKG---NVDLVFLFDGMSLQDFQKILDFPMKDMVKKLSNTSYQFAAV 196

RESULT 9

US-60-487-610-1788
; Sequence 1788, Application US/60487610

QY 175 QYSEBFIHFTKPEFQNNPNRSLKIPITQLLGRHTTATGLRKVVVRELFNITNGARKNAF 234
Db 197 QFSTSYKTEFDSDYVKKWDPDALLKHVKHMLLLTNTFGAINYVATEVFREELGARPAT 256
QY 235 KILFLLTGBEGDPLGYEDVPELDREGVIRYIVIGVGDAPRSEKSRQELNVTASKPRPD 294
Db 257 KVLIIITDGE--ATDSGNIDAAD-----IIRYIIGIKHFTKESQSTLHKFASKPASE 309
QY 295 HVFOJNNFEALKTIQNLREKIFAIEGTQTQSSSFEHEMSOEGFSAAITSNGPLLSTVG 354
Db 310 FVKILDTKEKLDLFTLELOKLIYIEGHSKDLTSFNMELSSGISADLSRHGAVGVAG 369
QY 355 SYDNAGGVF--LYTSKEKSTFTNTRVDSMDMDAYLGAAA--IILNRVQSLVGLGAPRYQH 412
Db 370 AKDWAGGFLDKADLQDDTTFIGNBPLTPEVRAGYLVTVTWLPSRKQKTSLLASGAPRYQH 429
QY 413 IGLVAMFR--QNTGWESNANVKTQICAGYFASGLSCSYVDVDSNGSTDLVLGAPHYEQT 470
Db 430 MGRVLLFOBPQGGHWSQVQTHGTQIGSYFGELCGVDVDQGETELLIGAPLYFGEQ 489
QY 471 RGGQVSVCPPLPRGORARWOCDAV--LYGEOGQPMGRFGAALTVLGDVNGDKLTQVAIGAP 528
Db 490 RGRVFIY-----QRRQLGFEEVSELOQDGPYLPORFGEAITALTIDNGDLVDVAVGAP 544
QY 529 GEEDNRGAVLYLPHGTSGSGISPSHSQRIAGSKLSPRLQYFQCSLSSGGQDLTMDGLVDLTV 588
Db 545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEGTVQVLSGIQMFGRSIHGKVDLEDGLADVAV 601
QY 589 GAQGHVLLLRSQPVLRVKAIMEFNPVARNVFCNDQV--KGKEAGEVRVCLHVQKSTR 647
Db 602 GAESQMIVLSSRPVYDMVTLMSFSPAETPVHEVECSYSTSNKMEGVNITTCFQI--KSLY 660
QY 648 DRLREGQISQVTVYDLDALDSGRPHSRVAFNFKNSTREOTQVLGLTQTCETLKLQLPNCI 707
Db 661 PQF--QGRUVANLTYTLQDDGHRTRRRGLFPGGRHELRRNIAT--TSMSCDTPSFHPFVCV 718
QY 708 EDPVSPVLRNLFSL---VGTPLS--AFGN-----LRPVLAEADAORLFTALPFPFKNCGN 757
Db 719 QDLISPINVSNFSLWEEBGTDRQRAQKIDPILRPSLHSETWEI-----PFERKNCGE 773
QY 758 DNIQODDLSITFSFMSLDCLVGGPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVS 817
Db 774 DKKEANLRVSFSPARSRALRLTAFASLSVELSLNLEEDAYVWQLDLHFPGLSPRKVE 833
QY 818 TLQORSORSWLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTVNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCCELPESRLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887
QY 876 LGNKLKLLKANVTSENN---MPTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNTFASEN 931
Db 888 WGSVELHANVTCNNEDSLLLEDNSATTI---IPILYPINILIOQOEDSTLYVSFTPKGP 944
QY 932 TSRVMQHOYQV---SNLQORSPL-ISLVFLVPLRLNQTIVWRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKMYQVRIQPSIHNDIPTLEAVGVGPQPPSEGPITHQWSVQMEPPV--PCHVED 1002
QY 987 -ERLPKSHSD--FLAELRKAPVNCISAVCQRIQCDIPFGIOEBFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAAEPCLPGALFRCPVW-----FRQELVQVIGTLELVGEIE 1044
QY 1044 TSHNHLIVSTAELFNDSVFTLLPGOGAFVRSOTETKVEPEVNPPLPLIYVSSVGGILL 1103
Db 1045 AS-SWFSLCSSLSISFNSSKHFLYGSNASL-AQVMKVDDVYVYKQMLYLYVLSGIGILL 1102
QY 1104 LLALITAALYKLGFFKRYQKDMMSSEG-GPPGAEP 1136
Db 1103 LLLLIFIVLYKGVFFKRNLEKMEAGRGVNGIP 1136

; GENERAL INFORMATION:

```

; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1788
; LENGTH: 413
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-60-487-610-1788

```

Query Match 19.4%; Score 1166; DB 7; Length 413;

Best Local Similarity 57.8%; Pred. No. 2.9e-60;

Matches 227; Conservative 61; Mismatches 103; Indels 2; Gaps 1;

```

Qy 1 FNLDTENAMTFOENARFGSGVQLOGSRVVGAPQEIIVAAANORGSLYQCSTGSCPEI 60
Db 20 FNLDTTELTAFRVDSDAGSGSVQVYANSWVVGAPQITAAQTGGYQCGYSTGACEPI 79
Qy 61 RLQVPEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOPK 120
Db 80 GLQVPEAVNMSGLSLASTTSPQLLACGPTVHHECGRNWYLTGLCLLGLPT-QLTOR 137
Qy 121 FPEALRCQPEDSDIAFLIDGSGIIPHDPRMKWSTVNEQLKSKTFLSLMQVSEEF 180
Db 138 LPVSRQBCPRQEQDVLIDGSGISSRNFAFMNFVRAVISQFQESTQSLMQFSNKF 197
Qy 181 RIHFTPEFONNPNRSLIKPITOLLGRTHATGLRVVRELFNITNGARKNAFKILFL 240
Db 198 QTHFTPEFRSSNPLSLASVHQLQGTTFATAIQNVHRLPHASTGAARDATKILIVI 257
Qy 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKRGQELNVTASKPRDHRVQIN 300
Db 258 TDGKKEGSLDYKDVPMDAAGIIRVAGVGLAFQNRNSWKLNDIASKPSQEHIFKVE 317
Qy 301 NFEALKTIONLRKEIFAEGTQTGSSSSSEHEHMSQSGFSAATTSNGLPLSTVGSYDNAG 360
Db 318 DFDALKDIONLRKEIFAEGTQTGSSSSSEHEHMSQSGFSAVFTPDGVLGAVGFTWSG 377
Qy 361 GVLFLYSKESKSTRNMTVRDSDMNDAYLGAAA 393
Db 378 GAFLYPPNMSPTFINMSQENVDMDRSTYLGPPSA 410

```

RESULT 10

```

; US-60-487-610-2494
; Sequence 2494, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2494
; LENGTH: 1179
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-60-487-610-2494

```

Query Match 19.4%; Score 1144; DB 7; Length 1179;

Best Local Similarity 29.0%; Pred. No. 1.7e-58;

Matches 341; Conservative 213; Mismatches 451; Indels 172; Gaps 39;

```

Qy 45 GSLYQCDYSTGS--CEPI-RLQVP-----VBVAVNMSGLSLAATTPPQLLACGPTVHQ 95
Db 65 GPLRCSLVODBIILCHPVEHVPIPKGRHGVTVVRSHHGVLICI-----QVLVRRP--HS 117
Qy 96 TCSNTYVKGCLFGLSGNLRQOPQ-----SDIAFLIDGSGIIPHDPRMKWSTVMEQL-- 119
Db 118 LSSELT---GTCSLLGPDRLRPOQANFFDLENLLDPDARVDTGDCYSNKEGGEDDVNTA 174
Qy 120 KPPEALRGCPQED-----SDIAFLIDGSGIIPHDPRMKWSTVMEQL-- 164
Db 175 RRRALKEEEDKEEEDDEEBEAGTEIAIILDGSGIDPPDFQAKDFINMWRNFYE 234
Qy 165 KKSCTLFLSMQYSEEFRIHFTKBFQNNPNRSLIKPITQLLGRTHATGLRKVYRELFN 224
Db 235 KCPECNFALVQYGGVIQTDFDLRDSQDVMASTARVQNIITGVSVTKTASAMQHVLDSIF 294
Qy 225 INGARKNAPKILFLITDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKSRBEL 284
Db 295 SSHGSRKASKVMVVLTDGGIFEDPLNLTIVINSKMQGVFERFAIGVGBEFKSARTAREL 354
Qy 285 NTVASKPPRDHVDVQINNFEALKTIONOLREKIFAIEGTOTGSSSSSEHEHMSQEGFSAIT 344
Db 355 NLIASDPDETHAFKVTNYMALDGLLSKLYNISHEGT---VGDHALYQLAQIGFSAQIL 411
Qy 345 SNGP-LLSTVSGSYDAGGVFLY-TSKEKSTFINMTRVDSMDNDA-----YLGAAAAILRN 398
Db 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAATAAADAEEAAQYSLGVAVVLHKT 471
Qy 399 RVQSLVGLAPRYQHIGLVAMFR-QNTGMWESNANV-KGTQIGAYFCASLCSVDVDSNGST 456
Db 472 CSLSYVAGAPRYKHHG--AVFELQKEGRBASFLPVLEGEQMGSYFGSELCPCVDIDMGST 529
Qy 457 DLVLICAPHYEOTRGQGVSVCPPLRGQARWQCDALVYGEQOPMGRFGAALTVLGDVN 516
Db 530 DFLVLAAPYHVHGEGRVYVYRLSE-QDGSFSLARILSGHPGFTWARFGFANAAGDLS 588
Qy 517 GDKLTDVAIGAP---GEEDNR--GAVLYFHTGSGSIGSPSHSQRIAGSKLSPLRYFGQ 570
Db 589 QDKLTDVAIGAPLEGGFADGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
Qy 571 SLSSGGQDLTMDGLVDLTGCAQGHVLLRSQPLRVKAIWEPNPREVARNVFECNDQVVK 630
Db 648 SMAGSFDISGDGLADITVGLQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
Qy 631 KEAGEVRVCLHVQKSTRDLREGIOISVVYTDALDSGRPHSRSAVNETKNSTRRTQVL 690
Db 698 NGVVNVLCFEI-SSVTTASEGLREALNFLTLDVDVGQRRLQCSVRSCLGCLREWS 756
Qy 691 GLTQTCETLKLQLPN---CIEDPVSPVILRNFLNFSLVGTPLSAFAGNLRPLVLAEDAORLFT 746
Db 757 SGSQLCEDL-LLMPTGELCEEDCFNSAVKVSQYL-QTPEGQTDHPQILDRYTEPPAI 814
Qy 747 ALFPPEKNCNDNICQDDLSITPFSNLDCLVWGGPRENVTVTVRNDGEDSVRTQVTF 806
Db 815 FQLPYEKACKNKLFCVAELQLA-TTVSQDELVVGLTKELTLNLTNSGSDSYMTSMALN 873
Qy 807 FPLDLSYRVKVTSLQNSQRSWRLACESASSTEVSGALKSTCSINHPIPPENSEVTFNI 866
Db 874 YPRNLO-----LKRQKPPSPNIQCDPQPV---ASVLMNCRIQHPVL-KRSSAHVS 923
Qy 867 TFDVDSKASLGNKLLKANVTSENN---MPRTNKTEFQ---LELPVKYAVYVMTVSHGV 919
Db 924 VMQLEENAFNRTADITVTVNSNERWSLANETHTLQFRHGFVAVLSPKPSIMYVNTGQGL 983
Qy 920 S---TKVINFNTASENTSRVMOHQOVNLSGORSPLISLVFLVPLVRLNQTVMWDRPQTFSE 977
Db 984 SHKBEFLFHVHGEN---LFGAEYO-----LQICVPTKRLGLQVAVVKKLTRTQ 1028
Qy 978 NLSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIPFFGQIEBFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVVEHWSVSCVIA-----SDKENVTVAEIS 1073

```



```
QY 1038 FDWYIKTSHNHLIVST-----AEILFNDSVFTLLPGQAFVRSOTETKVEPEVNPPL 1091
Db 1074 WD-----HSEELKDVTELIQILGEISFNKSLYEGLNAEH--RTKITVVFLLKDEKYHSL 1125
QY 1092 PLVIGSSVGGLLIALITALAALYKLGFFKROYKOWMSE 1128
Db 1126 PIIKGSVGGLLVILVILFKCGFFRKRYQOOLNLE 1162

RESULT 11
PCT-US03-18234-2
; Sequence 2, Application PC/TUS0318234
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: Antibodies that bind alphaE Integrin
; FILE REFERENCE: 1855-2025002
; CURRENT APPLICATION NUMBER: PCT/US03/18234
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 10/173,551
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(18)
PCT-US03-18234-2

Query Match 19.4%; Score 1140; DB 1; Length 1179;
Best Local Similarity 28.9%; Pred. No. 2.8e-58;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSYLQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTPSPOLLACGPTVHQ 95
Db 65 GPLHRCSLVQDEILCHVEHPVPIKGRHGVTVVRSHHGLICI-----QVLVRP--HS 117
QY 96 TCSNTVVKGLCLFLGSLNLQPOQ----- 119
Db 118 LSSELT---GTCSLGPDLPQAQANFFDENLDDPARVDTGDCYNKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLDGGSGIIPHPFRMKWVSTWMEQL-- 164
Db 175 RQRALEKEEEDKEEEDBEAEAGTEIAIILDGSGSIDPPQFQAKOFISNMNRNFE 234
QY 165 KKSKTFLSLMQYSEEPRIHFTFEFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFN 224
Db 235 KCFECNPAVQYGVIGVQTEFDLRDSQDVMASLARVQNIQVGSVTKTASAMQHVLSIFT 294
QY 225 ITNGARKNARKIILLTDGKFGDPLOYEDVPELDREGVIRVVGDAFRSEKSRQEL 284
Db 295 SSHGSRKASKVMVLTDGIFEDPLDITVINSPKMQYVERFAIGVGEFKAARTAREL 354
QY 285 NTVASKRPRDHVQINNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQSGFAAIT 344
Db 355 NLIASDDPEHAFKVTNYMALDGLLSKLRNIIISMEGT---VGDALHYLQAQIGFSAQIL 411
QY 345 SNGP-LSTVGSYDWMAGGVFLY-TSKEKSTFINNTRVDSIDMND-----YLGAAAAIILRN 398
Db 412 DERQVLLGAVGAFDWSGGLLYDTRSRGRFLNQTAADAAAEAAQVSYLGYAVAVLHKT 471
QY 399 RVQSLVLGAPRYOHIGLVAMFR-QNTQWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
Db 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFLPVLEGEQMGSGYFSGELCPVDIDMDGST 529
QY 457 DLVLIGAPHYVEQTRGOVSVCPILPRGORARWOCDAVLVYCEQCPWGRFCAALTVLGDVN 516
Db 530 DFLVRAAPFVTHGEGRGVTVYRLSE-QDGSFSLARILSLGHPGFTNARFGFAAAMGDLS 588
QY 517 GDKLTDVAIGAP-----GEEDNR--GAVYLPFGTSGSGISPSHSQRIAGSKLSPRLQYFGQ 570
```

```
Db 589 QDKLTDAVGAIPLEGFGADDGASFGSVITYNG-HMDGLSASPQRIRASTVAPGLQYFGM 647
QY 571 SLSSGGDLTMDGLVDLTGCAQGHVLLRSQPLVRVKAIMEFNEPREVARNVFECNDQVVKG 630
Db 648 SMAGGFDISDGLADITVGTGQAVVFRSRPVVRLKVSMAFTPSALP-----IGP 697
QY 631 KEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDIALDSGRPHSRVAVFETKNSTRRTQVVL 690
Db 698 NGVVNVRLCFEI--SSVTTASESGLEALLNFTLDVVGKQRRRLQCSVRSCLGLREWS 756
QY 691 GLTOTCETLKLQLPN-----CIEDPVSPVIRLNFSLVGTPLSAFAGNLRPVLAEDAORLFT 746
Db 757 SGSQLCEDL-LLMPTEGELCEDCFSNASVKVSYQL-QTPEGQTDHPQILDRTYTFPAI 814
QY 747 ALFPPFEKCGNDNICODDISITFSFMSLDCLVVGGPREFNVTVTVDNDEGDSVRTQVTF 806
Db 815 FQUPYBKACNKLFCVAELQLA-TTVSQOELVVGLTKELTNLNLNSGEDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQNRORSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNI 866
Db 874 YPRNLQ-----LKRQKPPSPNIQDDPPQV---ASVLIMNCRIGHPLV-KRSSAHVSV 923
QY 867 TPDVDSKASLGNKLLKANVTSENN-----MPRTNKTEPQ---LELPVKYAVVMVVTSHGV 919
Db 924 VMQLEENAFNRTADITVTVTNSNERRSLANETHTLQFRHGFVAVLKSPSIMYVNTQGL 983
QY 920 S--TKYLNFTASENTRVMQHOYQVSNLQGRSLPISLVLPVRLNQTVINDRPOVTFSE 977
Db 984 SHHKEFLFHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVNCSTIAVCQRIQCDIPFFGIEEFNATLKNLS 1037
Db 1029 ASVTCTWSQERACAYSS-VQHVVEHWSVCVTA-----SDKENVTVAAEIS 1073
QY 1038 FDWYIKTSHNHLIVST-----AEILFNDSVFTLLPGQAFVRSOTETKVEPEVNPPL 1091
Db 1074 WD-----HSEELKDVTELIQILGEISFNKSLYEGLNAEH--RTKITVVFLLKDEKYHSL 1125
QY 1092 PLVIGSSVGGLLIALITALAALYKLGFFKROYKOWMSE 1128
Db 1126 PIIKGSVGGLLVILVILFKCGFFRKRYQOOLNLE 1162

RESULT 12
US-60-487-610-1507
; Sequence 1507, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Honglin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS, CURRENT APPLICATION NUMBER: US/60/487,610
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1507
; LENGTH: 1149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1507

Query Match 18.8%; Score 1103.5; DB 7; Length 1149;
Best Local Similarity 27.7%; Pred. No. 3.5e-56;
Matches 343; Conservative 212; Mismatches 485; Indels 197; Gaps 44;

QY 1 ENLDTENAMTFQENARG-EGQSVVOL---QGSRVVVGAQOEIVAAANORGLYOCDDYSTGS 56
Db 1 FNVDDVKNMTSGFVEDMFGYTVQQYENEEGKWLIGSLVPLVGPKRKTGDVYKCPVGRGE 60
QY 57 CEP-IRLQVPVEA-----VNMSLGLSLAATTPSPOLLACGPTVHTQCSNTVVKGL 106
```



```
Db      61  SLPCVKLDLPVNTSIPNTEVKENMTFGSTL-VTNPNGGFLACGPLYAYRGCHLHYTTGI 119
Qy      107  CFLFGSNLRQOPKPEALRCPOEDSDIAFLIDGSGSIIPHDFRMKEMVSTWMEQLK- 165
Db      120  CSDVSPTFQVNSIAP--VQECSTQ-LDVIIVLDGNSIYPWD--SVTAFNLNLLERNDI 174
Qy      166  -KSKTLFSLMOYSEBFRIHFTFKFQNNPNRSLIKPITQLLGR--THATGLRKVVRBLF 223
Db      175  GPKQTQGVIVQYGENVTHERNLNKSSTEELVAAKKIVQGGQTMTALGIDTARKEAF 234
Qy      224  NITGARKNAKFIKLLTDGSKFGDPLGYEDVPELDEGVIRVIVGVGDAPR-----SE 278
Db      235  TEARGARRGVKVMVIVTDGESH-DNHLKKVIOQDENIQRFSAIILGNSRGNLSTE 293
Qy      279  KSRQELNVTASKPRDRHVQINNFEALATIQNLREKIFAIEGTQGTSSSFEHEMSQEG 338
Db      294  KPVEEIKSIASEPTEKHFNFVSDLEALVTIVKTGLERIFALEATADQSAASFEMMSQTG 353
Qy      339  FSAAITNSGPLLSTVSGYDAGGVPLYTSKE-----KSTF-INMTRVDSMDNDAYLYAA 392
Db      354  FSAHYSQDWMVGAVDWMGTVMQKASQIIIPRNTTFVESTKNEPL-ASYLGTVV 412
Qy      393  AILLRNVQSL-VLGAPRYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSDVD 451
Db      413  NSATASSGDVLYIAGQPRYNTGTQVLIYRMEDGNIKILQTLSGEIQSYFGSILTTDID 472
Qy      452  SNGSTDVLILGAPHY-----YEQTR-GGOVSVCPPLPRQORARWQCDVLY 495
Db      473  KDSNTDILLVAGAPMTGTEKEEQKVVYVALNQTREYOMSLPIKQTCSSRQHNSCTT 532
Qy      496  GEOGQPMWG-RFGAALTVLVDGNGDKLTDVAIGAPGEEDNRGAVLPHGTSGSGSPSHSQ 554
Db      533  ENKNEPCGARGFATAAANKDLNLDGFNDIVIGAPLEDHGGAVIYHG-SKTIKRYEVAQ 591
Qy      555  RIAGSKLPRLOYFGOSLGGODLTWQGLVDLTVGAQGHVLLRSQPLRVLRKAIEMFNPR 614
Db      592  RIPSQGDGKTLKFFQGSIHGMDLNGDGLTDVTIGLGAALFWSRDAVAVKVTMNFEPN 651
Qy      615  EVARNVFECNQVVGKQAG--EVRVCLHVO-KSTRDLRREGQIQSVTVYDLADSGRPH 671
Db      652  KVINIQKNCH--MEGKETVCINATVCFDVKLSKEDTIYEADLQ----YRTVLDLSLQI 704
Qy      672  SRAVENET-----KNSTRRTQVLGLTQTCTELKLQLPNCIEDPVSPVILRNLFSLVGT 725
Db      705  SRSFSGTQERKQVQNRITVRKSE-----CTKHSFYMLHFDQDSVR---ITLDENLI-D 753
Qy      726  PLISAFGNLRPVLAEDAORLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLLVGGPRE- 784
Db      754  PENG-----PVLDDSLPNSVHEIIPFAKDCGNKEKICISDLJHVAATTEKDLIIVRSDNK 808
Qy      785  FNVTYVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSORSWELACESASTSVSGAL 844
Db      809  FNVSLTVKNTKDSANTRTIHVHSPNLVFSGLEAIOKD-----SCESN----- 851
Qy      845  KSTSCSINHPIIPENSEVTFTITFDVDSKASLGN-KLLKANVTSENMPRTNKTETFOLE 903
Db      852  HNICTKVGPFLRGEMVTFKLFPQNTSYLMENVTIYLSATSDSEEPETLSDNVNIS 911
Qy      904  LPKVAVVMVTVTSHGVSTKYLNTASNTSRVMQHOYQVSN-----LGORS-----L 950
Db      912  IPKVVEVGLQFYS-SASEVHISIAANETVPEINSTEIDIGNEINIFYLIRKSGSPMPPEL 970
Qy      951  PLSLVF-----LVPVRLNQTIVDRPQVTFSENLSSTCHTKE-----RLPS 991
Db      971  KLSISFPNMTSNGYPVLYPTGLSS-----SENANCRPHIFDPFINSNGKMTT 1019
Qy      992  HSDFLAELRKAPVNCSTAVCORIOCDIPFFGQIE-----EFNATLK 1033
Db      1020  STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKTFFIKSYFSSLNLTIR 1076
Qy      1034  GNLSEFDWYIKTSHNLLIVSTAEILFNDSVFTLLPQCGAFVRSQETKVPPEVFNPLPL 1093
```

```
Db      1077  GEL-----RSNASLVLJSSN-----OKRELAIQISKOGLPGRVPL 1112
Qy      1094  --IVGSSVGLLLALITAAALYKLGFFKROYKDMWSE 1128
Db      1113  WVILLSAFAGLLLLMLLLALWKGFFKRLPKKKMEK 1149

RESULT 13
US-10-474-794-250
; Sequence 250, Application US/10474794
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-794-250
```

Query Match 18.6%; Score 1093.5; DB 6; Length 1179;
Best Local Similarity 27.8%; Pred. No. 1.4e-55;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

```
Qy      1  FNLDTEVAMTFQENARG-FGQSVVQL---OGRSVVVCAPQEIIVANORGSLYQCDYSTGS 56
Db      29  FNVVDKNSMTFGSPVEDMFGYTVQOYENEEGKWVLIGSLVGPQKNRTGDYKCPVGRGE 88
Qy      57  CEP-IRLQVPEA-----VNMSLGLSLAATTSPPQLLACQPTVHQTCSENTYVKGL 106
Db      89  SLPCVKLDLPVNTSIPNTEVKENMTFGSTL-VTNPNGGFLACGPLYAYRGCHLHYTTGI 147
Qy      107  CFLFGSNLRQOPKPEALRCPOEDSDIAFLIDGSGSIIPHDFRMKEMVSTWMEQLK 166
Db      148  CSDVSPTFQVNSIAP--VQECSTQ-LDVIIVLDGNSIYPWD-----VTAFLNLLK 198
Qy      167  -----SKTLFSLMOYSEBFRIHFTFKFQNNPNRSLIKPITQLLGR--THATGLRKVV 219
Db      199  RMDIGPKQTQGVIVQYGENVTHERNLNKSSTEELVAAKKIVQGGQTMTALGTD TAR 258
Qy      220  RELFNITNGARKNAKFIKLLTDGCKFGDPLGYEDVPELDEGVIRVIVGVGDAPR--- 276
Db      259  KEATFARGARRGVKVMVIVTDGESH-DNHLKKVIOQDENIQRFSAIILGNSYRGN 317
Qy      277  --SEKSRQELNVTASKPRDRHVQINNFEALATIQNLREKIFAIEGTQGTSSSFEHEM 334
Db      318  LSTEFVEEIKSIASEPTEKHFNFVSDLEALVTIVKTGLERIFALEATADQSAASFEMEM 377
Qy      335  SQEGFSAAITNSGPLLSTVSGYDAGGVPLYTSKE-----KSTF-INMTRVDSMDNDAYL 388
Db      378  SQTGFSAHYSQDWMVGAVDWMGTVMQKASQIIIPRNTTFVESTKNEPL-ASYL 436
Qy      389  GYAAAIILNRVQSL-VLGAPRYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCS 447
Db      437  GTVNSATSSGDVLYIAGQPRYNTGTQVLIYRMEDGNIKILQTLSGEIQSYFGSILTT 496
Qy      448  VDVSNGSTDVLILGAPHY-----YEQTR-GGOVSVCPPLPRQORARWQCD 491
Db      497  TDIDKDSNTDILLVAGAPMTGTEKEEQKVVYVALNQTREYOMSLPIKQTCSSRQHN 556
Qy      492  AVLVEGQOPMWG-RFGAALTVLVDGNGDKLTDVAIGAPGEEDNRGAVLPHGTSGSGISP 550
```

Db	557	SCCTENKNEPCGARGFTAI	AAVKDNLNDGNDIVIGAPLEDDHGGAVYIHG-SGKTIRK	615
Qy	551	SHSQRIAGSKLSRLQYFGQSLSGQDILTMDGLVDLTGAGQHVLLRLSRQVLRVKAIME	610	
Db	616	EYAQRIPSGGDKTKLFFQSGIHGEMDNGDGLDVTIGLGAALFWSRDVAVKVTWN	675	
Qy	611	FNPREVARNVFECNDQVVKKEAG--EVRVCLHVQ-KSTRDRLREGOIQSVVYDIALDS	667	
Db	676	FEPNKVNIQKNCH---MEGETVCINATVCFEVLKSKEDTIYEADLQ----YRVTLDS	728	
Qy	668	GRPHSRVAFNET-----KNSTRQTOVLGLTOTCETLKLQLPNCIEDPSPVLRLNFS	721	
Db	729	LROISRSFFSGTOERKQVRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFN	780	
Qy	722	LVTGTPLSAFGNLPRVLAEDAQRLFTALFPPEKNCNDNICQDDLSITTFMSLDCLVVG	781	
Db	781	LT-DPENG-----PVLDDSLPNSVHEIYIPFAKDCGNKERCISDLSLHVATTEKDLIVRS	834	
Qy	782	PRE-FNVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACSASSTEV	840	
Db	835	QNDKFNVSILTVKNTKDSAYNTRTIVHYSPLNLFVSGIEAIQKD-----SCSN----	881	
Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGN-KLLKANVTSENMPRTNKTE	899	
Db	882	-----HNITCKVGFPLRRGBMVTFKILFQNTSYLMENVTIYLSATSDSEPPETLSDNV	937	
Qy	900	FOLELPVKYAVYVMTSHGYSYKVLNFTASENTSRVMOHQYQVSN-----LQORS----	949	
Db	938	VNISIPKVEVGLQFYS-SASEYHISIAANETVEVINSTEDIGNINIFYLIRKSGSFP	996	
Qy	950	----LPISLVF-----LVPRLNQTVINDRPQVTFSENLSSTCHTKE-----	987	
Db	997	MPELKLISGFPNMTSNGYPVLYPTGLSS-----SENANCRPHIFEDPFSINSGK	1045	
Qy	988	RLPSHSDFLAELKAPVWNCISAVCQRIQCDIPFGIOE-----EPN	1029	
Db	1046	KMTTSTD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILKWKPTFIKSYFSSLN	1102	
Qy	1030	ATLKGNLSPDWYIKTSHNHLILVSTAEILFNDSVFTLLPGQAFVRSQETKVPPEVFN	1089	
Db	1103	LTRIGEL-----RSENASLVSSN-----QKRELAIQISKDGLPG	1138	
Qy	1090	PLPL--IVGSSVGLLLALITAAALYKLGFFKQYKQMMSE	1128	
Db	1139	RVPLWILLSAPAGLLMLLMLLALWKGIFGFKPLKKOMEK	1179	
RESULT 14				
US-10-474-794-307				
; Sequence 307, Application US/10474794				
; GENERAL INFORMATION:				
; APPLICANT: Carson-Walter, Eleanor				
; APPLICANT: St. Croix, Brad				
; APPLICANT: Vogelstein, Bert				
; APPLICANT: Kinzler, Kenneth				
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS				
; FILE REFERENCE: 1107.00179				
; CURRENT APPLICATION NUMBER: US/10/474,794				
; CURRENT FILING DATE: 2003-10-14				
; PRIOR APPLICATION NUMBER: 60/282,850				
; PRIOR FILING DATE: 2001-04-11				
; PRIOR APPLICATION NUMBER: 60/308,829				
; PRIOR FILING DATE: 2001-08-01				
; NUMBER OF SEQ ID NOS: 359				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 307				
; LENGTH: 1180				
; TYPE: PRT				
; ORGANISM: Rat				
US-10-474-794-307				
Query Match 17.9%; Score 1055; DB 6; Length 1180;				
Best Local Similarity 27.4%; Pred. No. 2.3e-53;				

Matches	346;	Conservative	196;	Mismatches	475;	Indels	246;	Gaps	48;
Qy	1	FNLDTENAMTFOENARG-FGQSVVQL--	QGSRVVVVGAPQEIIVAAANQRGSLYQCDYSTGS	56					
Db	29	FNVDVKNSSAFSGPVEDMEGYTVQYENEEGKWLIGSLVGPQKARTGDVYKCPVGR	88						
Qy	57	CSP-IRLVQVDEA-----VNSLGLSLAAATSPOLLACGPTVHOTCSENTYVKGL	106						
Db	89	AMPCKVLDLPVNTSIPNVTETKENMTFGSTL-VTNPNGGFLACGPLYAVRCGHLHYTTGI	147						
Qy	107	CFLFGSNLRQOQKPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKEMVSTM----	EQ 163						
Db	148	CSDVSPTFOVNSFAP--VQECSTQ-LDIVIVLDGNSIYP-----WESVIAPLNDL	196						
Qy	164	LKK-----SKTLFSLMOYSEERIHFTEKFEQNNPNRSLIKPITOLLG-RHTATGLRK	217						
Db	197	LKRMIDGPKQTQGVIGQYGENVTHEFNLNKYSSTEELVAANKI GRQGLQMTALGIOT	256						
Qy	218	VVRELFNITNGARKNAFKILFLTQGEKFGDPLGYEDVJPELDRBQVIRYIVGVGDAFR-	276						
Db	257	ARKEAFTARGARRGVKVMVITDGESH-DNYRLKQVTDCEDENIQRFSTAILGHYNR	315						
Qy	277	----SEKSRQELNTVASKPPRDHVFIQINNFEALKTIQNLREKI PAIECTQTGSSSSFEH	332						
Db	316	GMLSTEKFYEETKSTASEPTEKHFNVSDDELALVTIVKALGERIFALEATADQSAASFEM	375						
Qy	333	EMSQGFSAAITNSGPLLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDND----	AY 387						
Db	376	EMSQGFSAHYSQDWMLGAVGAYDNQVTVWQKANQMVIPHNTTFTQTEPAKNEPLASY	435						
Qy	388	LYAAAILLRNVQSLVGLAPRYQHIGLVAMFRQNTGMMESNANVKGTOIGAYFGASLCS	447						
Db	436	LGTVNSATIPGDVLYIAGQPRYNTGQVVIYKMGEDGNINILQTLGGEIGSYFGSVLT	495						
Qy	448	VDVDSNGSDVLVIGAPHY-----YEOTR-GQOVSVCLPLRQORARWQCD	491						
Db	496	IDIDKSDYDILLVGLAPMYMTGTEKEQGVYVYVYVQTRFYOMSLEPIRQTCSSLKON	555						
Qy	492	AVLYGEOGQPMG-RFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVLYPHGTSGGISP	550						
Db	556	SCTKENKNEPCARFGTAIAVKDLNVDFNDVIGAPLEDDHAGAVYIYHG-SGKTIRE	614						
Qy	551	SHSQRIAGSKLSPLQYFGQSLSGGODLTMDGLVDLTGAGQHVLLRLSRQVLRVKAIME	610						
Db	615	AYAQRIPSGGDKTKLKFFQSGIHGEMDNGDGLTDVTIGLGAALFWARDVAVKVTWN	674						
Qy	611	FNPREVARNVFECNDQVVKKEAG--EVRVCLHVQ-KSTRDRLRBQOIQSVVYDIALDS	667						
Db	675	FEPNKVNIQKNCR--VEGKETVCINATMCFHVKLSKEDSIYEADLQ----	YRVTLDS 727						
Qy	668	GRPHSRVFNET-----KNSTRROTQVLGLTOTCETLKLQLPNCI-----	EDPVS 712						
Db	728	LQIISRFSFGTQERKIQRNITVRESE-----CIRHFSYMLDKHDFQD	770						
Qy	713	PVLRNLNFSLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCNDNICQDDLSITFSFM	772						
Db	771	SVRVTLDFNLT-DPENG-----PVLDDALPNSVHEHI PPAKDCGNKERCISDLTLNVST	824						
Qy	773	SLDCLVVGQPRE-FNVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLA	831						
Db	825	EKSLIIVKSOHDKFNVSILTVKNGKDSAYNTRTVQHSPLNLI FSGIEIQD-----S	876						
Qy	832	CESASSTSVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLL-LKANVTSEN	890						
Db	877	CESN-----QNITCRVGFPLRAGETVTKIIFQNTSHLSNAILHLSATSDESSE	927						
Qy	891	NMPRTNKTEFOLELPVKYAV----YMWVTSHGYST-----KYLNTFASENTSRVMOHQ	939						
Db	928	PLESLNDNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNINVF	985						
Qy	940	YQVSNLQGRSLP-----ISLVF-----LVPRLNQTVIND-----RP-----	971						
Db	986	YTIRKRGHPMPPELQLSISFPNLDTADGYPVLYPIG-----WSSSDNVNCRPSRLEDPFG	1039						

```
Qy 972 -----QVTF-----ENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIQ 1016
Db 1040 INSGKWTISKSEVLKRGTIQDCSTC-----GVATITCSLLPSLSQ 1082
Qy 1017 CDI-----PFGIOEBSF---NATLKNLSPDWIKTSHNHLIIVSTABILFNDSVFTLL 1067
Db 1083 VNVSLMLWKPTF-IRAHFSSNLTLRGELK-----SENSLTLSSN----- 1123
Qy 1068 PQGAFVRSOTETKVEPEFVNPLPL--IVGSSVGGGLLLALITAALYKLGFFKQYKDM 1125
Db 1124 -----RKRELAIQISKDGLPGRVPLWVILLUSAFAGLLMLLILALWKIGFFKRLPKK 1177
Qy 1126 MSE 1128
Db 1178 MEK 1180

RESULT 15
US-60-493-369-8
; Sequence 8, Application US/60493369
; GENERAL INFORMATION:
; APPLICANT: Bruce DOMON
; APPLICANT: Tao HE
; APPLICANT: Xiaolong ZHANG
; APPLICANT: Karen KETCHUM
; APPLICANT: Ian McCAFFERY
; TITLE OF INVENTION: PANCREATIC DISEASES TARGETS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001476
; CURRENT APPLICATION NUMBER: US/60/493,369
; CURRENT FILING DATE: 2003-08-08
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-493-369-8

Query Match 17.9%; Score 1054; DB 7; Length 1179;
Best Local Similarity 26.7%; Pred. NO. 2.6e-53;
Matches 326; Conservative 219; Mismatches 494; Indels 182; Gaps 44;

Qy 1 FNLOTENAMTEQ-ENARCFGSUVQL---QCSRVVVGAPQEIIVANQSGSYQC--DYST 54
Db 30 YNVLPEAKIFSGPSSEQFGYAVQOFINPKGNWLLVSGPWSGFPENRMGDIYKCPVDLST 89
Qy 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTVYKGLC 107
Db 90 ATCEKLNQTSIPNVTETMKNLSGLILTRNMGTCGFLTCGPLWAOCCGNQYTTGVC 149
Qy 108 FLFGSNLRQPKPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEWSTVMQK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVWCDSENSIYPWD--AVKNFLEKFVQGLDIG 205
Qy 166 KSKTLFSLMQYSEERPIHFTKFPQNNPNRSLKIPITQLLG-RTHATGRLKVVURELN 224
Db 206 PTKTOVGLIOYANNPRVVFNLTNYKTEEMIVATSQTSQYGGDLTNTFGAIQYARKYAS 265
Qy 225 ITNGARKNAFKILFLTDGKFGDPLGYEDVIPELDREGVIRYIGV-----GDAFRSEK 279
Db 266 AASGRRSATKVMVVVTGDGESH-DGSMKAVIDQCNDHNIIRFGIAVLGYLNRNALDTKN 324
Qy 280 SRQELNTVASKPPRDRHVQINNFEALKTIONLREKIFAIEGTOTGSSSFEHMSQEGF 339
Db 325 LIKEIKATASIPTEYFFNVSDAALKEKAGLGEQIFSIEGTVQG-GDNFQMEMSQVGF 383
Qy 340 SAAITSNGP--LLSTVGSYDAGGVFLYTSKSTFINMT--RVDSMDN-DAYLGYAAAI 394
Db 384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSYILGSVAA 443
```

Search completed: November 25, 2003, 14:38:49
Job time : 17.9635 secs

```
Qy 395 ILNRVQSLVLAGAPRYOHIGLIVAMFRONTGMESNANV-----KGTQIGAYFASLCSV 448
Db 444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIOAHRGDQGSYFSGVLCV 498
Qy 449 DVDNSGSTDLVLIGAPHYEQTR--GGVSVCPPLPRQARWOCDAVLGECQOPWREG 506
Db 499 DVDKOTITDVLVAGPMYMSDLKKBEGRVYLFITKEGILGQHO---FLEGPBGIENTRFG 555
Qy 507 AALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KLSPR 564
Db 556 SAATAALSDINMDGFNDVIVGSPLENQNSGAVIYNGHQT-IRTKYSQKLTGSDGAFRSH 614
Qy 565 LQYFGQSLGGQDLTWDGLVDLTVGACGHVLLRSQPVLRVKAIMEFNPREFARNVFECH 624
Db 615 LQYFGRSLDGYGLNGDSITDVSIGAFGVQVQLWSQSIADVAIEASTPEKI--TLVNKN 672
Qy 625 DOVVKGKEAGEVAVCLHVQKSTRDLREGIOISVTVTYDLALD-----SGRPHSRAVNETK 680
Db 673 AQII-----LKLCF-----SAKPRPTKQNNQVAIVNITLDADGFSRSTSRGLPKNN 721
Qy 681 NSTRROTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRNLNLSLVGTPLSAFGLNRPVLA 738
Db 722 ERCLQKNVNVQAQSCPEHIYIQEPS---DVVNSLDLVDLSLENPCTS-----PALE 772
Qy 739 EDAQRLFTALFPPEKNCGNINICODLSITF---SPMSLDCVLVGGPREFNVTVTRND 794
Db 773 AYSETAKVFSIPPHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKKN 832
Qy 795 GEDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRLACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTVEVTCVAAASQKSVACDVGY 880
Qy 854 PIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFQLELPLVKYAYMV 913
Db 881 PALKREQOVTTFINFDENLQ-NLQNASLSFQALSSEQEBENKADNLVNLKIPLLYDAEIH 939
Qy 914 VTSHGVS TKYLNFTASENTS RVNMQHYQ-----VSNLQORSLPISLVFLV----- 958
Db 940 LT-RSTNINFEISSDGNVPSIV-HSFEDVGPKFIFSLKVGSPVSMATVIIHIPQYTK 997
Qy 959 -----PVRLNQTWIDRPOVTP-SENLSSTCHTKERLPSH 992
Db 998 KNPLMYLTGVQTDKAGDISCNADINPLKIGQT-----SSVSFKSENFR---HTKE----- 1045
Qy 993 SDFLAELRKAPVWNCIAVCORIQCOIPFFGIOEENFATLKGNSLPDWYIKTSHNHLIV 1052
Db 1046 -----LNCRTASCNSVTCWLKDVHMKGEYFVNVVTRINWGTFPASSTFTQVLT 1093
Qy 1053 STABI-LFNDSVFTLLPGQAGFVRSOTETKVEPEFVNPL-----LP--LIVGSSVGGLL 1104
Db 1094 AAAEINTYNEIYVI-----EDNTVTIPLMIMKPKDEKAEVPTGVIIGSIAGILL 1143
Qy 1105 LALITAAALYKLGFFKQYKDM 1125
Db 1144 LLALVAILWKLGFKRYEKH 1164
```

Result No.	Query			ID	Description
	Score	Match	Length		
1	5852	99.5	1153	1	US-08-173-497-3
2	5852	99.5	1153	1	US-08-286-889-3
3	5852	99.5	1153	1	US-08-485-618-3
4	5852	99.5	1153	1	US-08-362-652-3
5	5852	99.5	1153	2	US-08-605-672-3
6	5852	99.5	1153	2	US-08-482-293A-3
7	5852	99.5	1153	2	US-08-943-363-3
8	5852	99.5	1153	3	US-09-193-043-3
9	5852	99.5	1153	4	US-09-688-307A-3
10	5821.5	98.9	1152	2	US-08-476-062A-43
11	5821.5	98.9	1152	5	PCR-US96-01314-43
12	5821.5	98.9	1152	6	5424399-2
13	3459	58.8	1163	2	US-08-476-062A-44
14	3459	58.8	1163	5	PCR-US96-01314-44
15	3436	58.4	1163	1	US-08-173-497-4
16	3436	58.4	1163	1	US-08-286-889-4
17	3436	58.4	1163	1	US-08-485-618-4
18	3436	58.4	1163	1	US-08-362-652-4
19	3436	58.4	1163	2	US-08-605-672-4
20	3436	58.4	1163	2	US-08-482-293A-4
21	3436	58.4	1163	2	US-08-943-363-4
22	3436	58.4	1163	3	US-09-193-043-4
23	3436	58.4	1163	4	US-09-688-307A-4
24	3401	57.8	1161	1	US-08-173-497-2
25	3401	57.8	1161	1	US-08-286-889-2
26	3401	57.8	1161	1	US-08-485-618-2
27	3401	57.8	1161	1	US-08-362-652-2

QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVRELFINITNGARKNAFKILVLL 240
DB 197 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVRELFINITNGARKNAFKILVLL 256
QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTCSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTCSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGRGQVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGRGQVSVCLP 496
QY 481 PRGQARWQCDVAVLGEQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
DB 497 PRGQARWQCDVAVLGEQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRYKAIEMFNPVARNVFCNDQVVKGEAGVRVCHLVOKSTRDLREGQIQSVVT 660
DB 617 PVLRYKAIEMFNPVARNVFCNDQVVKGEAGVRVCHLVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRAVNETKSTRQTVGLGTQTCETLKLQPCIEDPVPSPVLRNLF 720
DB 677 YDLALDGRPHSRAVNETKSTRQTVGLGTQTCETLKLQPCIEDPVPSPVLRNLF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLETFALPFFKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLETFALPFFKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVTRNDGEDSVRTQVTFPFLDLSYRVKSTLQNSORSWRMLACESASSTEV 840
DB 797 GPREFNVTVTRNDGEDSVRTQVTFPFLDLSYRVKSTLQNSORSWRMLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMTVSHGVSTKYLNFTASENTSRVQHOYQVSNLQORSLPLSLVFLVPV 960
DB 917 QLELPVKYAVVMTVSHGVSTKYLNFTASENTSRVQHOYQVSNLQORSLPLSLVFLVPV 976
QY 961 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRQCDIP 1020
DB 977 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRQCDIP 1036
QY 1021 FFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFENDSVFTLLPQOGAFVRSQTEF 1080
DB 1037 FFGIQBEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFENDSVFTLLPQOGAFVRSQTEF 1096
QY 1081 KVEPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137
DB 1097 KVEPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMWSEGGPPGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCCEPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCCEPI 76
QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVRELFINITNGARKNAFKILVLL 240
DB 197 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVRELFINITNGARKNAFKILVLL 256
QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTCSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTCSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

```
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLF 540
DB 497 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSRLOQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSRLOQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVENETKNSRROTQVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
DB 677 YDLALDSGRPHSRAVENETKNSRROTQVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSAFNGLRPVLAEDAQLFTALPFPEKNCNDNIQDDLSITFSMSLCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEDAQLFTALPFPEKNCNDNIQDDLSITFSMSLCLVVG 796
QY 781 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDESDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANTSENNWPRNTKTEF 900
DB 857 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANTSENNWPRNTKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLSPLSLVLPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLSPLSLVLPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCNSIAVCQRIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCNSIAVCQRIQCDIP 1036
QY 1021 FFGIQEENATLKNLSFDWYIKTSHNHLIVSTABILFNDVSTLPGGAFVRSQTET 1080
DB 1037 FFGIQEENATLKNLSFDWYIKTSHNHLIVSTABILFNDVSTLPGGAFVRSQTET 1096
QY 1081 KVEPPEVPNPLPLIVGSSVGLLLALITAAALYKLGFFKRYKDMMSBEGGPPGABPQ 1137
DB 1097 KVEPPEVPNPLPLIVGSSVGLLLALITAAALYKLGFFKRYKDMMSBEGGPPGABPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVKGCLFGLFGSNLRQOPK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVKGCLFGLFGSNLRQOPK 136
QY 121 FPEARLGCPEQSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 FPEARLGCPEQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEXFGDPLGYEDVIPLEDRGVIRYVIGVGDAFRSEKSRQELNTVASKPRDRHVFOIN 300
DB 257 TDGEXFGDPLGYEDVIPLEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPRDRHVFOIN 316
QY 301 NFEALKTQNLREKIFAIEGTQTGSSSFEHEMSEGFSAAITSGNPILLSTVGSYDWAQ 360
DB 317 NFEALKTQNLREKIFAIEGTQTGSSSFEHEMSEGFSAAITSGNPILLSTVGSYDWAQ 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLF 540
DB 497 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSRLOQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
```

Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616
Qy 601 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLNF 720
Db 677 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSVRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDCGDSVRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 916
Qy 901 QLELPVKAVYVMVTSHGVSFTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKAVYVMVTSHGVSFTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNOTVIVDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCVSIACVQRIQCDDIP 1020
Db 977 RLNOTVIVDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCVSIACVQRIQCDDIP 1036
Qy 1021 FFGIOEERNATLKGNSLFDWYIKTSHNHLLIIVSTAEILFNDSVFTLLPGQCAFVRASQTET 1080
Db 1037 FFGIOEERNATLKGNSLFDWYIKTSHNHLLIIVSTAEILFNDSVFTLLPGQCAFVRASQTET 1096
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPPQ 1137
Db 1097 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPPQ 1153

RESULT 4

US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Beet Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQBIIVANQSGSLYQCDSYSGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQBIIVANQSGSLYQCDSYSGSCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCTSENTYVKGCLCFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCTSENTYVKGCLCFPGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDCGSGSIIIPHDFFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDCGSGSIIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHTFFKEFQNNPNRSLIKPITOLLGRTHATGLRKVYRELFINTGARKNAFKILPVL 240
Db 197 RIHTFFKEFQNNPNRSLIKPITOLLGRTHATGLRKVYRELFINTGARKNAFKILPVL 256
Qy 241 TDGKFKGDPGLGYEDVPELDREGVIRVIVGVGDFAFRSEKSKQELINTVASKPRDHVFCQIN 300
Db 257 TDGKFKGDPGLGYEDVPELDREGVIRVIVGVGDFAFRSEKSKQELINTVASKPRDHVFCQIN 316
Qy 301 NFEALKTIQNLREKIPIAIEGTQTGSSSSFEHMSQGFSAATISNGPILLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIPIAIEGTQTGSSSSFEHMSQGFSAATISNGPILLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTREGQVSVCPPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTREGQVSVCPPL 496
Qy 481 PRGORARWQCDAVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNDRGAVVLF 540
Db 497 PRGORARWQCDAVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNDRGAVVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616
Qy 601 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLNF 720
Db 677 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSVRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840

Db 797 GPREFNVTVTRNDGEDSVRTQVTFPPDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPPVKYAVYVMTSHGVSTKYLNFTASENTSRVQHQQVQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPPVKYAVYVMTSHGVSTKYLNFTASENTSRVQHQQVQVSNLQORSLSPLSLVLPV 976
Qy 961 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQDIP 1020
Db 977 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQDIP 1036
Qy 1021 FFGIOBEFNATLKNLSFDMYIKTSHNHLIVSTAELFNDSVFTLLPGOGAFVRSQTE 1080
Db 1037 FFGIOBEFNATLKNLSFDMYIKTSHNHLIVSTAELFNDSVFTLLPGOGAFVRSQTE 1096
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVVGAPQEI VAAANQORSLLYQCDYSTGSCDEP 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVVGAPQEI VAAANQORSLLYQCDYSTGSCDEP 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPVTHOTCSNTYVKGICFLFGLNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPVTHOTCSNTYVKGICFLFGLNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEMVSTVMEQLKKSKTFLSLMQYSEBF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEMVSTVMEQLKKSKTFLSLMQYSEBF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKVVRELFNITNGARKNAKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKVVRELFNITNGARKNAKILFL 256
Qy 241 TDGEXFGDPLGYEDVI PELDREGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVFQIN 300
Db 257 TDGEXFGDPLGYEDVI PEADREGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVFQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSEFHEMSEGEFSAATITNGPLLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSEFHEMSEGEFSAATITNGPLLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQOVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQOVSCPL 496
Qy 481 PRQQRARWQCDALYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 497 PRQQRARWQCDALYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDLTVAQAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDLTVAQAQGHVLLRSQ 616
Qy 601 PVLRYKAIMENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
Db 617 PVLRYKAIMENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVSPVLRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVSPVLRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPFEKNCNDNICQDDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPFEKNCNDNICQDDLSITFSFMSLCLVVG 796
Qy 781 GPREFNVTVTRNDGEDSVRTQVTFPPDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVTRNDGEDSVRTQVTFPPDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPPVKYAVYVMTSHGVSTKYLNFTASENTSRVQHQQVQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPPVKYAVYVMTSHGVSTKYLNFTASENTSRVQHQQVQVSNLQORSLSPLSLVLPV 976
Qy 961 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQDIP 1020
Db 977 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQDIP 1036

QY 1021 FFGIOEERNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQGGAFVRSQTET 1080
Db 1037 FFGIOEERNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQGGAFVRSQTET 1096
QY 1081 KVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPPGABPO 1137
Db 1097 KVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPPGABPO 1153
RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3
Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches: 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQORGLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSNTYVVKGLCFGLFSGNLROQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSNTYVVKGLCFGLFSGNLROQPOK 136
QY 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 180

Db 137 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHTTFKEFQNNPRLSIKPIITOLGRTHATGLRKVRELFINITGARKNAFKILPFL 240
Db 197 RIHTTFKEFQNNPRLSIKPIITOLGRTHATGLRKVRELFINITGARKNAFKILPFL 256
QY 241 TDGKFGDPLGYEDVIPLEDRGVIRVYVIGVDGAFRSEKSRQELNITASKPPRHVFOVN 300
Db 257 TDGKFGDPLGYEDVIPLEDRGVIRVYVIGVDGAFRSEKSRQELNITASKPPRHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIECTQTGSSSSPFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIECTQTGSSSSPFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 480
Db 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 496
QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDAIGAIGEDNDRGAVLYF 540
Db 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDAIGAIGEDNDRGAVLYF 556
QY 541 HGTSGSISPSHSORIASKLSPLQYFGQSLSGQDITMDGLVDLTGAGQHVLLLRQ 600
Db 557 HGTSGSISPSHSORIASKLSPLQYFGQSLSGQDITMDGLVDLTGAGQHVLLLRQ 616
QY 601 PVLRYKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRYKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSCRPHSRVAFNETKSTRTOVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720
Db 677 YDLALDSCRPHSRVAFNETKSTRTOVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRULTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRULTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANNVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANNVTSENNMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSKYNFTASNTSRVMQHQYOVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSKYNFTASNTSRVMQHQYOVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCTKRLPSHSDFLAELRKAPVNCVNCIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCTKRLPSHSDFLAELRKAPVNCVNCIAVCQRIQCDIP 1036
QY 1021 FFGIOEERNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQGGAFVRSQTET 1080
Db 1037 FFGIOEERNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQGGAFVRSQTET 1096
QY 1081 KVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPPGABPO 1137
Db 1097 KVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPPGABPO 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.58; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.28; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGCEPI	60
Db	17	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGCEPI	76
Qy	61	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHTQCSNTYVKGCLFGLGSLNRQOPQK	120
Db	77	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHTQCSNTYVKGCLFGLGSLNRQOPQK	136
Qy	121	FPALRCPCQEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMEQLKSKTFLSLMQYSEEF	180
Db	137	FPALRCPCQEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMEQLKSKTFLSLMQYSEEF	196
Qy	181	RIHFTKFQNNPNRSLKIPITQLGRTHATGLRKVVRELFINITNGARKNAFKILFLL	240
Db	197	RIHFTKFQNNPNRSLKIPITQLGRTHATGLRKVVRELFINITNGARKNAFKILVVI	256
Qy	241	TGKEKFDPLGYEDVIPEDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	300
Db	257	TGKEKFDPLGYEDVIPEDREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	316
Qy	301	NFALKTIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG	360
Db	317	NFALKTIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG	376

Qy	361	GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAAAIILNRNVQSLVLCAPRYQHIGLVAMER	420
Db	377	GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAAAIILNRNVQSLVLCAPRYQHIGLVAMER	436
Qy	421	QNTGMWESNANVKGTQIGAYFGASCLSDVDNSGSTDVLIGAPHYYEOTRGQSVVCPL	480
Db	437	QNTGMWESNANVKGTQIGAYFGASCLSDVDNSGSTDVLIGAPHYYEOTRGQSVVCPL	496
Qy	481	PRGORARWOCDAVLYGEOGQPMGRFGAALTVLVDGVDNGDKLTDVAICAPGEEDNRGAVYLF	540
Db	497	PRGORARWOCDAVLYGEOGQPMGRFGAALTVLVDGVDNGDKLTDVAICAPGEEDNRGAVYLF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGVAQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGVAQGHVLLRSQ	616
Qy	601	PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT	660
Db	617	PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT	676
Qy	661	YDLALDSGRPHSRVFNETKSTRROTQVLGLTQTCETIKLQLPNCIEDPVSPIVLRNF	720
Db	677	YDLALDSGRPHSRVFNETKSTRROTQVLGLTQTCETIKLQLPNCIEDPVSPIVLRNF	736
Qy	721	SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFPEKNCNDNICDDLSITFSFMSLDCLVWG	780
Db	737	SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFPEKNCNDNICDDLSITFSFMSLDCLVWG	796
Qy	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLONQORSQSWRLACSSASTEV	840
Db	797	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLONQORSQSWRLACSSASTEV	856
Qy	841	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASIGNKLLKANTYSENMPRTNKTEF	900
Db	857	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASIGNKLLKANTYSENMPRTNKTEF	916
Qy	901	QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSLFLVPV	960
Db	917	QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSLFLVPV	976
Qy	961	RLNQTVINDRPOQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1020
Db	977	RLNQTVINDRPOQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1036
Qy	1021	FFGIOEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSOTET	1080
Db	1037	FFGIOEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSOTET	1096
Qy	1081	KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGAEPO	1137
Db	1097	KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGAEPO	1153

RESULT 8

US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03

```
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEVSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEVSTVMEQLKSKTFLSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILPLL 240
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 316

Qy 301 NFEALKTQNLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILNRVOSLVILGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILNRVOSLVILGAPRYQHIGLVAMFR 436

Qy 421 QNTGWESNANYKGTQIGAYFGASLCSDVDNSNGSTDVLVIGAPHYEQTGGQVSVCP 480
Db 437 QNTGWESNANYKGTQIGAYFGASLCSDVDNSNGSTDVLVIGAPHYEQTGGQVSVCP 496

Qy 481 PRGQARWOCDAVLGEOGPGRFGAALTVLGDVNGDKLTDVAITGAPGEDNRCGAVLYF 540
Db 497 PRGQARWOCDAVLGEOGPGRFGAALTVLGDVNGDKLTDVAITGAPGEDNRCGAVLYF 556

Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLRSQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLRSQ 616

Qy 601 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

Qy 661 YDLALDSGRPHSAFNETKSTRQTVLGLTQTCETLKLQPCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSAFNETKSTRQTVLGLTQTCETLKLQPCIEDPVPVILRLNF 736

Qy 721 SLVGTPLSAFNLRLPVLAEADQRLFTALPFFKNGCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRLPVLAEADQRLFTALPFFKNGCNDNI CODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYKYSTLQNRQSRWRLACESASSTEV 856

Qy 841 SGALKSTCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANVTSNNNPRNTKTEF 900
Db 857 SGALKSTCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANVTSNNNPRNTKTEF 916

Qy 901 QLELPVKYAVYVMVTSKYLNFNTASNTSRVMQHQYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSKYLNFNTASNTSRVMQHQYQVSNLQORSPLISLVFLVPV 976

Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP 1036

Qy 1021 FFGIQBEFNATLKNLSFDMYIKTSHNHLIIVSTABIFLNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FFGIQBEFNATLKNLSFDMYIKTSHNHLIIVSTABIFLNDVSFTLLPGQAFVRSQTET 1096

Qy 1081 KVBPFVEPNPLPLIVSGSVGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1137
Db 1097 KVBPFVEPNPLPLIVSGSVGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688.307A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEVSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEVSTVMEQLKSKTFLSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILPLL 240
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 316

Qy 301 NFEALKTQNLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
```

Db 317 NFALKTIONQREKIPIAIEGTOTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGORARWOCDAVLVYGEQGPWGRFGAALTVDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLVYGEQGPWGRFGAALTVDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy 601 PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLVHKSTRDLRREGQIOSVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLVHKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKNSRTOVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAVNETKNSRTOVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFENLRPLVAEDAQRLLFTALPFEKXNCNDNICQDDLITFTFSMSLDCLVVG 780
Db 737 SLVGTPLSAFENLRPLVAEDAQRLLFTALPFEKXNCNDNICQDDLITFTFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLVSRYKSTLQNRQSQRSWRLACSSASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPDLVSRYKSTLQNRQSQRSWRLACSSASSTEV 856
Qy 841 SGALKTSKCSINHPISPENSEVFNITFDVDSKASLGNKLLKXANTSENMPRTNKTEF 900
Db 857 SGALKTSKCSINHPISPENSEVFNITFDVDSKASLGNKLLKXANTSENMPRTNKTEF 916
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVMQHOVQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVMQHOVQVSNLQORSIPISLVFLVPV 976
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQC 1020
Db 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQC 1036
Qy 1021 FFGIOBEFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096
Qy 1081 KVPEFVFNPLPLIVGSSVGLLLALITAAALKYKLGFFKROYKDMMSSEGGPPGAEPO 1137
Db 1097 KVPEFVFNPLPLIVGSSVGLLLALITAAALKYKLGFFKROYKDMMSSEGGPPGAEPO 1153

RESULT 10

US-08-476-062A-43
Sequence 43, Application US/08476062A

Patent No. 5877275

GENERAL INFORMATION:

APPLICANT: Arnaout, M. Amin

TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

Zip: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCES/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-476-062A-43

Query Match 98.9%; Score 5821.5; DB 2; Length 1152;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSLNRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSLNRQPOK 136
Qy 121 FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEMVSTVMEQLKKSKTLFSLMQYSEEF 180
Db 137 FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEMVSTVMEQLKKSKTLFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGURKVVRRELFTNGARKNAFKILVLI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGURKVVRRELFTNGARKNAFKILVLI 256
Qy 241 TDGEKFGDPLGYEDYIPELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFOIN 300
Db 257 TDGEKFGDPLGYEDYIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFOIN 316
Qy 301 NFEALKTIONQREKIPIAIEGTOTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDMAG 360
Db 317 NFEALKTIONQREKIPIAIEGTOTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGORARWOCDAVLVYGEQGPWGRFGAALTVDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLVYGEQGPWGRFGAALTVDVNGDKLTVAIGAPGEEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600

```

Db 556 HGTSGGISPSHSORIIAGSKLSPRLQVFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLAVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAVFNETKNSRRQTVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVFNETKNSRRQTVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 735
Qy 721 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVVMVTSHGVSSTKYNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTSHGVSSTKYNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNCVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNCVCSIAVCQRIQCDIP 1035
Qy 1021 PFGIOEBENATLKGNSFDWIKTSHNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1036 PFGIOEBENATLKGNSFDWIKTSHNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1095
Qy 1081 KVEPEVNPPLIIVGSSVGGLLIUALITAALYKLGPFKROYKDMWSSGGPPGAEPQ 1137
Db 1096 KVEPEVNPPLIIVGSSVGGLLIUALITAALYKLGPFKROYKDMWSSGGPPGAEPQ 1152

```

RESULT 11

```

PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:

```

```

; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston

```

```

; STATE: Massachusetts

```

```

; COUNTRY: U.S.A.

```

```

; ZIP: 02110-2804

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555x
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)

```

CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/380,167

```

```

; FILING DATE: 30-JAN-95

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: John W. Freeman

```

```

; REGISTRATION NUMBER: 29,066

```

```

; REFERENCE/DOCKET NUMBER: 00786/267001

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (617) 542-5070

```

```

; TELEFAX: (617) 542-8906

```

```

; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

```

```

Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

```

```

Qy 1 FNLDTENAMTFQSNARGFGQSVVQLQGSRRVVGAPQEIIVAAORGSLYQCDYSTGSGCEPI 60
Db 17 FNLDTENAMTFQSNARGFGQSVVQLQGSRRVVGAPQEIIVAAORGSLYQCDYSTGSGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWSTVMEQLKXSKTLFSLMQVSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKXSKTLFSLMQVSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITMGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITMGARKNAFKILVVI 256
Qy 241 TDEKFKGDPGLGYEDVIPEDLREGVIRYVIGVGDAFRSEKSRQELNVTASKPPRDHVPQIN 300
Db 257 TDEKFKGDPGLGYEDVIPEDLREGVIRYVIGVGDAFRSEKSRQELNVTASKPPRDHVPQIN 316
Qy 301 NFEALKTIQNLREKIPAIEGTQTGSSSSFEHEMSQGFSAAITNSGPNLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIPAIEGTQTGSSSSFEHEMSQGFSAAITNSGPNLLSTVGSYDWAG 376
Qy 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAATAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAATAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 496
Qy 481 PRQARWQCDVAVLYGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAAGEEDNRGAVYLP 540
Db 497 PRG-RARWQCDVAVLYGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAAGEEDNRGAVYLP 555
Qy 541 HGTSGSGISPSHSORIIAGSKLSPRLQVFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSORIIAGSKLSPRLQVFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAVFNETKNSRRQTVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVFNETKNSRRQTVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 735
Qy 721 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 915

```


Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
Qy 1021 PFGIOEBFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDSDVFTLLPGQAGAFVRSOTET 1080
Db 1036 PFGIOEBFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDSDVFTLLPGQAGAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2
; LENGTH: 1152
5424399-2

Query Match 98.9%; Score 5821.5; DB 6; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTQENARGFQSGVWOLQGSRRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFQSGVWOLQGSRRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQCSNTYVKGCLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQCSNTYVKGCLCFLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKXWSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKXWSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPRLSKIPITOLLGRTHATGLRKVRELFTNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPRLSKIPITOLLGRTHATGLRKVRELFTNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLGVEDVPELDREGVIRYVIGVDGDAFRSEKSRQELNTVASKPRDHVFOIN 300
Db 257 TDGKFGDPLGVEDVPELDREGVIRYVIGVDGDAFRSEKSRQELNTVASKPRDHVFOIN 316
Qy 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPFLSTVGSYDWAG 360
Db 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPFLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGVSVCPPL 480
Db 437 QNTGWESNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGVSVCPPL 496
Qy 481 PRGORARQCDAVLVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540

Db 497 PRG-RARWQCDAVLVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGQDGLTMDGLVDLTVGAGQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGQDGLTMDGLVDLTVGAGQGHVLLRSQ 615
Qy 601 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVRVCLVHVKQSTRDRREGIOQSVVT 660
Db 616 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVRVCLVHVKQSTRDRREGIOQSVVT 675
Qy 661 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 735
Qy 721 SLVGTPLSAFNGRLPVLAEADAQRLFTALFPFENKCGNDNICQDDLSITTFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGRLPVLAEADAQRLFTALFPFENKCGNDNICQDDLSITTFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSRWLACASSTEV 840
Db 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSRWLACASSTEV 855
Qy 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
Qy 1021 PFGIOEBFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDSDVFTLLPGQAGAFVRSOTET 1080
Db 1036 PFGIOEBFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDSDVFTLLPGQAGAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1152

RESULT 13
US-08-476-062A-44
; Sequence 44: Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

Qy	1	FNLDTENAMTPOENARGFQSVQVLOQSGRVVVGAPQBEIVAANQSGSIYQCDYSTGSCPEI	60
Db	20	FNLDEELTAFRVDSDAGFQSDWQYANSVWVVGAPQKITAAQNTGGLYQCYSTGACEPI	79
Qy	61	RLOVPVAVNMNLSGLSAAATSPSPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPOK	120
Db	80	GLQVPPBAVNMNLSGLSAASTSPSLACGPTVHHECGRNMYLTGLCFLGLPT--QUTQR	137
Qy	121	FPEALRGCPQBDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQLKTKGLTFLSMQYSEEF	180
Db	138	LPVSRQECPRQEQDIVFLIDGSGSISRNFAFMNFVRAVISQFORSTQFSLMQFNKPF	197
Qy	181	RIHFTTFKEFQNNPNRSUUKPIITOLLGRTHPTATGLRKVVBELFNITNGARKNAFKILFL	240
Db	198	QTHTEFEFRRTSNPLSLLASVHQLQGTYYTATAIQNVAHRLFHASYGARRDATKILIVI	257
Qy	241	TDGKFGDPLGYEDVIEPELOREGVIRYVIGVDAFRSEKSRQELANTVASKPRDRHVQIN	300
Db	258	TDGKKGESLDYKDVIPMDAAGIIRVAIGVGLAFQNRNSWKEUNDIASXPQBHIKPYE	317
Qy	301	NFEALKTIQONLREKIPAEICTQTQSSSSFEHEMSQEGFSAATISNGPILLSTVGSYDWAG	360
Db	318	DFDALKDOIQNLEKIPAEICTEITSSSSFELENAQEGFSAVFTPDGPVLGAVGSFTWSG	377
Qy	361	GVFLYTSKEKSTFTINMTRVJSDMNDAYLGAAAIILNRNVOSVLVGAPRYQHIGLVAMFR	420
Db	378	GAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVLGAPRYQHTGKAVIFT	437
Qy	421	QNTGMWESNANVKTOIGAVFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGGOVSVCPL	480
Db	438	QVSRQWRMAEVTQTOIGSVFGASLCSVDVDTDGSTDLLVLIGAPHYYEOTRGGOVSVCPL	497
Qy	481	PRGORARWQCDVAVLYGEGOPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF	540
Db	498	PRGWR-RWCDVAVLYGEGOPWGRFGAALTVLGVDVNGDKLTDVIGAPGEENRGAVYLF	556
Qy	541	HGTSGSISPSSHQRIAGSKLSPRLFQSGSLGGQDLTMDGLVDLTVGAGQHYLLLRSQ	600
Db	557	HGVLGSPISPSHSQRIAGSQLSRLRFQQAQSGQDLTDGLVDLAVGARGQVLLLRTR	616
Qy	601	PVLRAVKAIMENPREVARNVFECDNDVVKGKEAGEVRVCLHVQKSTDRLEREGOIQSVWT	660
Db	617	PVLWVGVMQMQIPAEIIPRAPECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676
Qy	661	YDLALDGRPHSRVAFNETKNSTRQTVLGLTCTCTLQLPNCIEDVPSPVLRLNF	720
Db	677	LDLALDPGRLSPRATFOETKNRSURVRVGLKAHCENFNLLLPSCVEDSVTPTTLRLNF	736
Qy	721	SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNGDNICQDDLSITTFSMJDLCLVVG	780
Db	737	TLVGKPLLAFLNRLPMLLAALQRYFTASLPFEKNCGADHICQDNLGIFSFPFGLKSLVG	796
Qy	781	GPREFNVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRKLACSSASSTEV	840
Db	797	SNLEINAEVMVWVNDGESYGTITTFSPAGLSYRYVAEGQKQGLRSLHLLTCDSPAPV--	854
Qy	841	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTFF	900
Db	855	SQGTWSTSCRINHILIFRGQAQITPLATFDVSPKAVLDGDRLLLTANVSENNTPTSKTTF	914
Qy	901	QLELPVKYAVVMVTVSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLQORSLPISLVLFP	959
Db	915	QLELPVKYAVVTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORLDPVSNFWVP	974
Qy	960	VRNLQTVMDRPOVTFSENLSSTCHTXERLPSHSDFLAELRKAPVNCSTAVCORICDI	1019
Db	975	VELNQEAVMMDVEVSHPONSLRCSSEKIPAPPASDFLAHQKPNVLDGSTAGCLRFPCDV	1034
Qy	1020	PFFGIGQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSQTE	1079
Db	1035	PSFSVQBELDFTLKNLSFGWVRQILOKKYSVSVSAEITFDTSVYSQLPQGEAFMRAQTT	1094
Qy	1080	TKPEPFEVPPNPLPIVGVSSVGGLLLLLALITAAALYKGLFFRKQYKDMNSE	1128

```

Db      1095 TVLEKYVHNPTPLIVGSSIGGLLLALITAVLYKVGFPRQYKEMWEE 1143
          | :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|
RESULT 15
US-08-173-497-4
; Sequence 4, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-4

Query Match           58.4%; Score 3436; DB 1; Length 1163;
Beat Local Similarity 60.6%; Pred. No. 1.le-287;
Matches 684; Conservative 148; Mismatches 291; Indels 6;

Qy    1  FNLDTENAMTFQENARGFGSVVOLGSRVVVGAPQEIIVAANGSLYOCYSTST
Db     20  FNLDTEELTAFRVDAGFGDSVVQYANSVVWGAPOKIIAANOIGGLYQCGYST
Qy    61  RLQVPVEAVNMSGLSLAATSPPPQLACGFTVHQTCSENTYYVKGLCFLPGSNL
Db     80  GLQVPPPEAVNMSGLSLASTSPSQQACGPTVHHEGRNMYLTGLCFLGLGT-
Qy   121  FPBALRCOPEDSDIAFLIDSGSIIPHDPRRKKEWVSTVMBOCLKSKTKTLFSLM
Db   138  LPVSQRQCPREQODIIFVLIDSGSISSRNPATMNFVRAVISOFQRPSTQFSLM
Qy   181  RIHTTRKEFNQNNPRLSPIKPIITQLLGRTHTATGLRKVVRELFNITNGARKNAF
Db   198  QTHETFPFEFRNTNPJSLASVHQLOGFTYTAIQNVHRLPHASYGAARDAI
Qy   241  TDGEKFQDPLGYEDVIPELDREGVIRVIYGVDARSEKSROBELNTVASKPPRD
Db   258  TDCKKESLDYKDVFPMADAAGIIRAIVGVGLAFQNRSWKELNDIASKPQOE
Qy   301  NFEALKTIQNOLREKIPIAIEGTOTGSSSPSEHEMSQEGFAAITNSGNPLLSTVGV
          : :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|

```

Db 318 DFDALDKIQNQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
Qy 361 GVFLYTSREKSTFINMTVRDSDMNDAYLGYAAAIILNRKQSVSLVGLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVDMRDSYLGYSTELALWKGVQSVLVLGAPRYOHIGKAVIFI 437
Qy 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTREGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYEQTREGGQVSVCP 497
Qy 481 PRQBARWQCDVAVLGEQOPWGRGGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEQHPWGRFGAALTVLGDVNGDKLTDVWIGAPGEBENRGAVYLF 556
Qy 541 HGTSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQCHVLLRSQ 600
Db 557 HGVLGPSISPSHSORIASQLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGOVLLRTR 616
Qy 601 PVLRYKATMEFNPRVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWVGVSQFIPABIPRSAFEQVWSEQTLVQSNICLYIDKRSKNLLGSRDIQSSVT 676
Qy 661 YDLALDSRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDALAPGRLSPRAFIOETKRSLSRVRVLGKAHCENFNLLPSCVEDSVIPIILRLNF 736
Qy 721 SLVGTPLGAFGNLRPLVAEDAORLFTALPPFEKNCNDNI CODDISITFSFMSLDCLVVG 780
Db 737 TLVGLPPLAFLNLRPMLAALQRYETASLPFEKNCADHI CODNLGISFSFPLKSLVVG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 SNLELNAEVMWMDGEDSYGTTITFSHPAGLSYRYVAEQKQQLRSLHLTC--CSAPVG 854
Qy 841 SCALKSTSCSINHPIFPENSEVTENITPDVDSKASLGNKLLILKANVTSENNMPRTNKTEF 900
Db 855 SQGTWSTSCRIINHLIFRGGAQITFLATFDVSPKAVGLDRKLLIANVSSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVYVMTVSHGYSTKYLNFAS-ENTSRVMOHQYQVSNLQORSPLISLVFLVP 959
Db 915 QLELPVKYAVYIVVSSHEQTKYLNFSSEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
Qy 960 VRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDI 1019
Db 975 VELNQEAVMMDVEVSHQPNSLURCSSEKIPASDPLAHIQKNPVLDCSIAGCLFRCDV 1034
Qy 1020 PFFGIOEENATLKNLSFDWIKTSHNHLIIVSTAELFNDVSFTLLPGQCAFVRSOTE 1079
Db 1035 PSFSVQOEELDFTLKGNLSFGWVRQILQKKVSVVSAEIIIDTSVYSQLPQGFAPMRAOTI 1094
Qy 1080 TKVEPPEVPNPLIVGSSVGLLLIALITALYKLGFFKROYKDMWSE 1128
Db 1095 TVLEKYKVHNPILIVGSSIGGLLLALITAVLYKVGFFKROYKEMWEE 1143

Search completed: November 25, 2003, 14:23:17
Job time : 17.5618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876
Sequence: 1 FNLDTEAMTFQENARGFGQ.....FKRQYKDMMSGGPPGAE PQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	5862	99.8	1153	1	RWHU1B cell surface glyco
2	4470	76.1	1153	2	S00551 leukocyte surface
3	3479	59.2	1163	1	RWHU1C cell surface glyco
4	1533.5	26.4	1170	2	S03308 cell surface glyco
5	1538.5	26.2	1163	2	F56126 lymphocyte fuction
6	1153	19.6	1179	2	A53213 integrin alpha-E c
7	1103.5	18.8	1151	2	A45226 integrin alpha-1 c
8	1085	18.5	1170	2	T45914 integrin alpha 2 s
9	1071	18.2	1178	2	S44142 VLA-2 protein homo
10	1068	18.2	1181	2	A33998 integrin alpha-2 c
11	1062	18.1	1180	2	A35854 integrin alpha-1 c
12	663	11.3	1039	2	A41131 lymphocyte-Peyer's
13	635	10.8	1038	2	S06046 integrin alpha-4 c
14	630	10.7	1035	2	T58409 integrin alpha-9 c
15	614.5	10.5	1041	2	T31437 integrin alpha cha
16	579.5	9.9	1054	2	JC7294 alphap integrin -
17	572.5	9.7	1051	2	A35761 cell surface glyco
18	567.5	9.7	1053	2	T55534 VLA-3 alpha subuni
19	555.5	9.5	1053	2	S44250 integrin alpha-5 c
20	542.5	9.2	1034	2	A36108 integrin alpha-v c
21	535	9.1	1044	2	T10050 integrin alpha-v c
22	532	9.1	1049	2	A20799 fibronectin recept
23	532	9.1	1073	2	B36429 integrin alpha-6 c
24	531.5	9.0	1072	2	A38457 integrin alpha-6 c
25	530.5	9.0	1051	2	A40021 integrin alpha-3 c
26	526	9.0	1048	2	A27421 integrin alpha-5 c
27	525.5	8.9	1091	2	A41543 integrin alpha-8 c
28	517	8.8	1044	2	T16516 integrin alpha-8 c
29	496.5	8.4	1146	2	S40311 integrin - fruit f

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C;Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; 152567

R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,

B.

A;Reference number: A31108; MUID:88315033; PMID:2457584

A;Accession: A31108

A;Molecule type: mRNA

A;Residues: 1-1153 <COR>

A;Cross-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148

A;Note: part of this sequence was confirmed by protein sequencing

R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A;Reference number: A28915; MUID:88257215; PMID:2454931

A;Accession: A28915

A;Molecule type: mRNA

A;Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA:

A;Note: the authors translated the codon TAC for residue 1129 as Thr

R;Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally ,

A;Reference number: A41600; MUID:92073318; PMID:1683702

A;Accession: A41600

A;Molecule type: DNA

A;Residues: 1-9 <SHE>

A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215

R;Arnaout, M.A.; Renold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A;Reference number: A94193; MUID:88190151; PMID:2833753

A;Accession: A30892

A;Molecule type: mRNA

A;Residues: 917-1042 <AR2>

A;Cross-references: GB:M18044

R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A;Reference number: A32218; MUID:89098893; PMID:2563162

A;Accession: A32218

A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>

A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FLE>
A:CROSS-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
A>Note: sequence extracted from NCBI backbone (NCBI:121963)
R:Pierce, M.W.; Renold-O'Donnell, E.; Todd III, R.F.; Arnsaut, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PIE>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: I52567; MUID:92144986; PMID:1346576
A:Accession: I52567
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:CROSS-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:CROSS-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802;881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVANQKSLYQCDYSTGSCRP 60
Db 17 FNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVANQKSLYQCDYSTGSCRP 76

Qy 61 RLQVPVAVNMSLGLSLAATSPPOLLAGCPVTHQCSNTYVKGCLFPGSNLRQKQK 120
Db 77 RLQVPVAVNMSLGLSLAATSPPOLLAGCPVTHQCSNTYVKGCLFPGSNLRQKQK 136

Qy 121 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIIVMBQAKSKTLFSLMOYSEF 180
Db 137 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIIVMBQAKSKTLFSLMOYSEF 196

Qy 181 RIHFTFKEFQNNPNRSLKPTQLLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
Db 197 RIHFTFKEFQNNPNRSLKPTQLLGRTHATGIRKVVRELFNITGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGVEDYIPVADREGVIRYVIGVDGAFRSEKSKOELNTVASKPRDRHVQFN 300
Db 257 TDGEKFGDPLGVEDYIPVADREGVIRYVIGVDGAFRSEKSKOELNTVASKPRDRHVQFN 316

Qy 301 NFPAKLTIONQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITSNGLLSTVGSYDWAG 360
Db 317 NFPAKLTIONQLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITSNGLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 496

Qy 481 PRGORARQWQDAVLYGSGQGPWGRFGAALTGLGVNGDKLTDVAIGAPEGEDNRGAYVLF 540
Db 497 PRGORARQWQDAVLYGSGQGPWGRFGAALTGLGVNGDKLTDVAIGAPEGEDNRGAYVLF 556

Qy 541 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAGHVLLLRSQ 616

Qy 601 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRREGQIQSVVT 660
Db 617 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPVLRINF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPVLRINF 736

Qy 721 SLVGTPLSAFGNRPVLAEDAQRLLFTALFPFKNCGNDNICQDDLSITFSPMSLDCLVVG 780
Db 737 SLVGTPLSAFGNRPVLAEDAQRLLFTALFPFKNCGNDNICQDDLSITFSPMSLDCLVVG 796

Qy 781 GPREFNVTVVRNMGDSYRTQVTFPPDLDSYRKVSTLQNRORSQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNMGDSYRTQVTFPPDLDSYRKVSTLQNRORSQSRWLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLQNLKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLQNLKLLKANVTSENMPRTNKTEF 916

Qy 901 QLELPVYAVYVMVYTSHGVSSTKYLNFTASENTSRVMQHVQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVYAVYVMVYTSHGVSSTKYLNFTASENTSRVMQHVQVSNLQORSLPISLVFLVPV 976

Qy 961 RLNOTVTDWRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNOTVTDWRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

Qy 1021 PFGIQEEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQGAFFVRSQTET 1080
Db 1037 PFGIQEEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQGAFFVRSQTET 1096

Qy 1081 KVEPFEPVNPPLILVSGSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1137
Db 1097 KVEPFEPVNPPLILVSGSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1153

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C:Accession: S00551; I59078

R:Pyte, R.

EMBO J. 7, 1371-1378, 1988

A:Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th

A:Reference number: S00551; MUID:88312584; PMID:3044779

A:Accession: S00551

A:Molecule type: DNA

A:Residues: 1-1153 <PVT>

A:CROSS-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983

A>Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: 159078; MUID:86287312; PMID:2942940
A:Accession: 159078
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;
Best Local Similarity 73.9%; Pred. No. 2.5e-302;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

QY 1 ENLDTENAMTFQENARGFGSVVQLOGSRVVVGAPOEIVAAORGSLYOCYSTGSCBPI 60
DB 17 ENLDTENAMTFQENARGFGSVVQLOGSRVVVGAPOEIVAAORGSLYOCYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATSPQQLACGPTVHQTCSNTYVYKGLCFGLFGSNLRQOPQK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVYKGLCFGLFGSNLRQOPQK 136
QY 121 FPAALRGCEQSDIAFLVDGSGSIIPHDFRRAKEFISTVMBOLKSKTLFSLMQVSEBF 180
DB 137 FPAALRGCEQSDIAFLVDGSGSIIPHDFRRAKEFISTVMBOLKSKTLFSLMQVSEBF 196
QY 181 RHFTPEKQNDNPRLKIPITOLLGRHTATGIRKVVRELPHNTNGARKNAFKILILI 240
DB 197 RHFTPEKQNDNPRLKIPITOLLGRHTATGIRKVVRELPHNTNGARKNAFKILILI 256
QY 241 TDEKEFGDPLGYEDVPEADREGVIRVIGVGADPAREKSRQBLNTVASKPRDHDVQFN 300
DB 257 TDEKEFGDPLGYEDVPEADREGVIRVIGVGADPAREKSRQBLNTVASKPRDHDVQFN 316
QY 301 NFALNTIONQLREKIFALEGTGTGSSSFEHEMSQEGFSAAITSNGPLSTVSGVDMAG 360
DB 317 NFALNTIONQLREKIFALEGTGTGSSSFEHEMSQEGFSAAITSNGPLSTVSGVDMAG 376
QY 361 GVPLYTSKESKSTFNNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHLGLVAMER 420
DB 377 GVPLYTSKESKSTFNNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHLGLVAMER 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLIGAPHYVEOTRGGOVSVCPL 480
DB 437 ENFGTWEPTHTSINGSOIGSYFGASLCSVDMDADGNTNLILIGAPHYVEOTRGGOVSVCPL 496
QY 481 PRGQARWQCDVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSGISPHSQRTAGSKLSPRLQYFGOSLGGQDLTDWGLVDLTVGAQHVLRLRQ 600
DB 556 YGASIASLASHSRRIIGAFHSPGLQYFGOSLGGQDLTDWGLVDLTVGAQHVLRLRQ 615
QY 601 PVLRVKAIMEFNPREVARNVFCNDQVWKGKEAGEVRVCLVHVKSTRDLRREGIOQSVVT 660
DB 616 PVLRVKAIMEFNPREVARNVFCNDQVWKGKEAGEVRVCLVHVKSTRDLRREGIOQSVVT 675
QY 661 YDLALDSGRPHSAFNETKNSTRROTQVLGLTQTCTETLKLQUNPCIEDPVSPIVLRLNF 720
DB 676 YDLALDSGRPHSAFNETKNSTRROTQVLGLTQTCTETLKLQUNPCIEDPVSPIVLRLNF 735
QY 721 SLVGTPLSAFNGLRPVLAEDAQRALTALPFPEKNCNDNIQQDLSITFSWLSLCLLVG 780
DB 736 TLVGEPLRSGNLRPVLAEDAQRALTALPFPEKNCNDNIQQDLSITFSWLSLCLLVG 795

RESULT 3

RWHUIC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A:Residues: 1-1163 <COR>

A:Note: this revision to the sequence from reference A35543 includes the carboxyl end

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecu

A:Reference number: A35543; MUID:90153906; PMID:2303426

A:Accession: A35543

A:Molecule type: DNA

A:Residues: 1-834 <CO2>

A:Note: this sequence has been revised in reference A36584

R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocy

A:Reference number: S00864; MUID:88166645; PMID:3327687

A:Accession: S00864

A:Molecule type: mRNA

A:Residues: 1-755, 'L', 757-1163 <CO3>

A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830

A:Note: part of this sequence was confirmed by protein sequencing

C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on

C:Genetics:

A:Gene: GDB:ITGAX; CD11C

A:Cross-references: GDB:119758; OMIM:151510

A:Map position: 16p11.2-16p11.2

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat ho

C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repe

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>

F:20-1107/Domain: extracellular #status predicted <EXT>

F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>

F:1108-1133/Domain: transmembrane #status predicted <TM>

F:1134-1163/Domain: intracellular #status predicted <INT>

F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status p

A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151

A;Experimental source: hepatoblastoma cell line HepG2
A;Note: sequence extracted from NCBI backbone (NCBI:P124326)
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.8%; Score 1103.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 4.3e-68; Indels 195; Gaps 44;
Matches 344; Conservative 212; Mismatches 486;

QY 1 FNLDTENAMTFQ-FCQSYVQL---QGSVVVVGAPQEIIVAAANQORGLSYOCDSYSGS 56
DB 1 FNVVKNMSTFSQVDMGTYVQYENEEGKWVIGSLVGPQPKRTGDVYKCPVGRGE 60
QY 57 CEP-IRLVQVPEA-----VMSLGLSLAATTPPOLLACGPTVHOTCSNTYVYKGL 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHLYTTGI 119
QY 107 CFLFGSNLRQOPKFEALRGCPQEDSDIAFLVDGSGSIIPHPFRKAEFISTVMEQLK- 165
DB 120 CSDVSPTFQVNVSIAP--VQECSTQ-LDIVIVLDGSGSIYPWDSVTA--FLNDLLKRMDI 174
QY 166 -KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLGR-THTATGIRKVVREL 223
DB 175 GPKQTQGVIVQYGENVTHEFNLNKYSSTEVLVAAKIVORGRTWTALGTDATKKEAF 234
QY 224 NITNGAKNAFKILILITDEKFGDPLGYEDVPIEADREGVIRYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKVMVITDGESH-DNHRLLKKVIOCDENIQRFSAIILGSYNRGNLSTE 293
QY 279 KSRQELNTVASKPRDHVFOINNFEALKTIONLRKIPAEIGTQTGSSSPHEMSQSG 338
DB 294 KFVEEIKSTASEPTKHFNFVSDALVTIVKTGLGERIFALEATADQSAAFEMMSQTG 353
QY 339 FSAAITSNGLLSTVSGYDMAGGVFLYTSKE-----KSTP-INWTRVDSMDNDAVLGYAA 392
DB 354 FSAHYQDWMLGAVGAYDNWGVVMQKASQIIIPRNTTFNVSTKXNEPL-ASYLGYYTV 412
QY 393 AIIILNRVOSL-VLGAAPRQVHIGLVAMFRONTQMWSNANVKQTQIGAFGSLCSVDVD 451
DB 413 NSATASSGDVLYTAGQPRVNHQTQVLIYRMEDCNILQTLSEQIGSYFGSILTTDID 472
QY 452 SNGSTDLVLIGAPHY-----YEOTR-GGQVSVCLPRGQRARQCDVAVLY 495
DB 473 KDSNTDILLVGAPMYMGTEKEEQKVYVVALNQTREYQMSLEPIKQTCSSRQHSCTT 532
QY 496 GEOGQPMW-RFGAALTVLGDVNGDKLTDVAIGAPGEDNDRGAVYLFHGTSGSGISPSHQ 554
DB 533 ENKNEPCGARFGTAIAAVDNLNDGNDIVIGAPLSDDHGGAVYIYHG-SGKTIKRYAQ 591
QY 555 RIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLVGAQGHVLLRSOPVLVRKAIMBNFR 614
DB 592 RIFSGDGGKTLKFFGSGIHGMDLNGDGLTDVIGLGGAAFPWSDVAVVYVMFEPN 651
QY 615 EVARNVECDNQVVKKEAG--EVRVCLHVQ-KSTRRLREGQIQSVVYVLDALDSGRPH 671
DB 652 KVIQKKNCH---MEKETVCINATVCFEVLKSKEDTIVEADLQ---YRVTLDSLRLQI 704
QY 672 SRAVFNET-----KSTRQTOVLGTQTCETLKLQLPNCIEDPVPVIRLNFSLVGT 725
DB 705 SRSFSGTQBRKVORNTIRKSEC-----TKHSFYMLDKRHFQDSVR---ITLDFNLT-D 755
QY 726 PLSAFNLRLPVLAEDAQRFTALFPPEKNGCNDNI CODDLSITFSFMSLCLVVGPGPRE- 784
DB 756 PENG-----FVLDDSLNPSVHEIYPPAKDCGNKKEKICISLDSLVHVAITEKDLLIVRQNDK 810
QY 785 FNVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQRWRLACBASSTEVSGAL 844
DB 811 FNVSLTVKNTKDSAYNTRTIVHYSPLNVFSGIEAIQKD-----SCSEN----- 853

845 KSTSCSINHPIPPENSEVTFNITFDVDSKASIGN-KLLLLKANVTSENNMPTNKTFFOLE 903
DB 854 HNITCKVGYPFLRRGEMVTFKILFQNTSYLSENVITVLSATSDSEEPETLSDNVNIS 913
QY 904 LPVKYAVYVYVTSHGSKYLNFTASENTSRVMQHQYQVSN-----LCQRS-----L 950
DB 914 IPVKYEVGLQFYS-SASEYHISIAANETVPEINSTEDEIGNIEINIFYLIRKSGSPMPEL 972
QY 951 PISLVF-----LVPVRLNQTVIWDPRQVTFSENLSSTCHTKE-----RLPS 991
DB 973 KLSISFPNMTSGYVPVLYPTGLSS-----SENANCRPHIFEDPFSINSKGKMTT 1021
QY 992 HSDFLAEIRKAPVNCVSIACQRIQCDIPFFQIQE-----EFNATLK 1033
DB 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFKSYFESSNLTR 1078
QY 1034 GNLSPDWYKTHSHNHLILVSTAEIILFNDSVFTLLPQOGAFVRSQRTETKVEPEVPNPL 1093
DB 1079 GEL-----RSENASLVLSN-----QKRELAIQISKDGLFGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKQYKDMWSE 1128
DB 1115 WVILLSAFAGLLMLLILALWKIGFFRPLKXKMEK 1151

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit
A;Reference number: A54402; MUID:94193847; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: GB:I25886; NID:9439695; PIDN:AAB59255.1; PID:9439696
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.5%; Score 1085; DB 2; Length 1170;
Best Local Similarity 27.6%; Pred. No. 8.6e-67;
Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFGQSYVQL---QGSVVVVGAPQEIIVAAANQORGLSYOC--DYST 54
DB 19 YNVGLPKAKIFSGPSSEQFGYAVQQFIPNKGWLLVGSWSPGPKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSNTYVYKGLC 107
DB 79 TTCEKLNLTQTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAOQCGSQYVYTTGVC 138
QY 108 FLFGSNLRQOPKFEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEQLK-- 165
DB 139 SDVSPDF-QLRSTFPAPVQTCF-SFIDVVVCDENSIYPWD--AVNLFLEKVFQGLDIG 194
QY 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLL-----GRTHATGIRKVVRE 221
DB 195 PIKTQMLQIYANPRVFNLTFRSKD---EMIKATSTQTFQYGGDLTNTFRAIOYARDT 251
QY 222 LFNITNGARKNAFKILILITDEKFGDPLGYEDVPIEADREGVIRYVIGV-----GDAFR 276
DB 252 AYSTAGGRPGATKVMVVVTDGESH-DGSKLKAVIDQCCKNINILRFGIAGLVGLYLRNALD 310
QY 277 SKSRQELNTVASKPRDHVFOINNFEALKTIONLRKIPAEIGTQTGSSSPHEMSQ 336
DB 311 TKNLKEIKAIASIPTEHFHFNVSDEADLEKAGTIGEQIFSIEGVQCG-GDNFQMEMSQ 369
QY 337 EGFSAAIT--SNGPLLLSTVSGYDMAGGVFLYTSKEKSTFTINMT--RVSDSMN-DAYLGYA 391

Db 370 VGFSAEYSPQNNILMLGAVGAYDWSGTGVVQKTPHGHILFSKQAEQILQDRNHSSYLGS 429
Qy 392 AAILLRNRVQSIVLGAPRYQHIGLVAMFRQNTGMWESNANV-----KGTOIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANTYTGQVLYSVN-----ENGNVTVQSQGGDQIGYFGSVL 484
Qy 446 CSVDVDSNGSDTLVLIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDAVLYGQGPWG 503
Db 485 CAVDVNKTITDVLVAGAPMYNDLKEEGRVYLTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTVLGVDNGDKLTDVAIGAPGEDNKGAVLPHGTSGSGISPSHSORIASG--KL 561
Db 542 RFGSAIALSDINMDGDFNDVVGSPLENQNSGAVIYNGHEGM-IRLAYSQKILGSDRAF 600
Qy 562 SPRLQYFGOSLGGDGLMDGLVLTGAGQGHVLLRQOPVLRKAIEMFNPREVARNVF 621
Db 601 SSHLYQFGRSLDGYDLNGDSITDVSQVAFQGVQVQVLSQSDADVSVDASFTPKXI--TLL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVVTYDLALD-----SGRPHSRVFN 677
Db 659 NKNAEI-----KLKLCF---SAKPRPTNQNNQVAIVYNTIDEDQFSRVSIRGLFK 707
Qy 678 ETKNSTRTQVGLGTQCE--TLKQLPNCIEDPVSPVILRLNPSL--VGTPLSAFQNL 733
Db 708 ENNERCLOKMTVSAQRCSEYIIHQEPS---DIISPLNLCMNISLENPGT-----756
Qy 734 RPLVAEDAQRFTALFPPEKXGNDNICODDLSITF---SFMSLDCLVWGPREFNVTV 789
Db 757 NPALAYSETVVFIPKHCQDGDGVCLISDLVNLVQQLPATQQOPPIVSNQNKLTFSV 816
Qy 790 TVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASST-EVSGALKSTS 848
Db 817 QLKNNKESAYNTEIVVDFSENLF-----ASWMPVDGTEVTCQIASQKSVT 864
Qy 849 CSINHIPIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFEQLPKY 908
Db 865 CNVGYPAKSKQOQVTFITNFDNLQ-NLQNAQASISFRALSESQEEENMADNSVNLKSLLY 923
Qy 909 AVYVMTVSHGVSTKYLNTFASNTSRVMQHOVQVSNLQOR-----SLPISLVFLV 958
Db 924 DAEIHIT-RSTNINYEVLSDGNVSVV-HSEF--DIGPKFISKVTGTVGPVUSMA---976
Qy 959 PVRLNQTVIMDRPQVTFSEN---LSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVIIHIPQVTKDNPLMYLTGVHTDQAGDISCEAEINPLKIGOTSSSVFSKSEN 1030
Qy 999 LRKAPVNCIAVCORIQCIDIFFGIOBEFNATLGNLSFDWIYKTSNHLIIVSTAEI- 1057
Db 1031 PRHIKELNCRATSCSNIMCWLDRDLQVGEYFLNVSTRIWNGTFAASTFTQVLTAAAEID 1090
Qy 1058 LPNDSVFTL-----LPGQGAFFVRSQVETETKVEPFE-VPNPLPLIVGSSVGGLLLALITA 1110
Db 1091 TYNPQIVYEENTVTP-----LTIKMPHEKVEVPTGVIVGSVIAGIILLALLVA 1140
Qy 1111 ALYKLGFFPKROYKDM 1125
Db 1141 ILWKLGFPPKRYEKM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A;Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;
Best Local Similarity 27.9%; Pred. No. 8.2e-66;

Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQOEIVAAANQSGSLYQC--DYST 54
Db 27 YNVGLPGAKIFSGPSSEQFGYSVQQLTNPQGNLLVGSWSPFENRMDGVYKCPVDLPT 86
Qy 55 GSECEPIRLQ-----VPVEAVNMSLGLSLAATSPOLLACGPTVHQTCESTYKGLC 107
Db 87 ATCEKLNQNSASISNVEIKTNMSLGLTLTRNPGTGGFLTCPLMAHQCGNYYATGIC 146
Qy 108 FLFGSNLRQOPQ---KFPEALRGCPQEDSDIAFLVDGSGSIIIPHDPRRAKEFISTVNEQL 164
Db 147 ----SDVSPDFQFLTSPAPVQACPSL-VDVVVVCDSESNLYP--WEAVNKLKVFVTGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLLG-RTHTATGIRKYVRE 221
Db 200 DIGPKKTQVALIQYANEPRIIFNLNDFETKEDVMVQATSETROHGGDLTNTFRAIEPARDY 259
Qy 222 LFNITNGARKNAFKILILITDGEKFGDPLGYEDVIEPADREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSGRPRGATKVMVVTGDESH-DGSKLTKTVIOQCNDDEILRFGIAVLGYLNRNALD 318
Qy 277 SEKSRQELNVTASKPPRDHVFQINNFEALQTIQNLREKIPAEIGTGTSSSSPEHEMSQ 336
Db 319 TKNLKEIKAIKASTPTERYFFNVADAEALLLEKAGTLGEQIFSEIGTVQG-GDNFQEMWAQ 377
Qy 337 EGESA--AITSNGPLLSVGSYDMAGGVFLYTSKEKSTFINMT--RYVSDMN-DAYLGYA 391
Db 378 VGFSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVLDQRNHSPLGYS 437
Qy 392 AAILNRVOSLVGAPRYOHIGLVAMFRQNTGMWESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAIISTEDGVHVFAGAPRANTYTGQVLYSVNK---QGNVTVIOSHRGDQIGSYFGSVLCS 494
Qy 448 VDVDNSGSTDVLVIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDAVLYGQGPWGFR 505
Db 495 VDVKDQITDVLVAGATPTMNDLKESGKVLFTTITKILNQHO---FLGPEGTGNARF 551
Qy 506 GAALTVLGVDNGDKLTDVAIGAPGEDNKGAVLPHGTSGSGISPSHSORIASGKLSPR- 564
Db 552 GSAIATLSDINMDGDFNDVVGSPVENENSGAVIYNGHQT-IRTKYSQKILGNGAFRR 610
Qy 565 -LQYFGOSLGGQDLTMDGLVLTGAGQGHVLLRQOPVLRKAIEMFNPREVARNVPEC 623
Db 611 HLQPFGRSLDGYDLNGDSITDVSIGALQVQLWSQSIADVAIEALFTP-----660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVVTYDLALD---SGRPHSRVFNET 679
Db 661 -DKITLLNKQAKITLKLCFRAEFPAGQNNQV---AILFMTLDDAGHSSRVTSGVFRN 717
Qy 680 KNSTRRTQVGLGTQTCET--LKLPLNCTIEDPVSPVILRLNPSLAFGNLRPVL 737
Db 718 SERFLQNMVNVVEQKCEHHISIQKPS---DVVNPLDLRVDISLENPGTS-----PAL 768
Qy 738 AEDAQRFTALFPPEKXGNDNICODDLSI-----TFSFMSLCLVVGGRPFNVT 788
Db 769 EAYSETVKVFSIPFVKECGSDGICISDLILDVQQLPATQTSF-----IVSNQNKRLTFS 823
Qy 789 VTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASST-EVSGALKST 847
Db 824 VILKNRGESANTVVLAEFSLENLFF-----ASFMPVDGTEVTECEVSSQKSV 871
Qy 848 SCSENHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET--FQLELP 905
Db 872 TCDVGYPAKSKQOQVTFITNFDNLQ-NLQNAQAINFOAFSESQ--ETNKADNSVSLTIP 928
Qy 906 VKYAVYVMTVSHGVSTKYLNTFASNTSRVMQHOVQVSNLQOR-----SLPISLV 955

```

Db 929 LLVDAEHLT-RSTNINFEISSDENAPSVIK-----SVEDIGPKFIFSLKVTAGSAPVSM 984
Qy 956 FLV-----PVRINOTVIMDRPQVTF-SENLS 980
Db 985 LVTHIQYTKENPLLYLTGIOTDQAGDISCTAEINPLKLPHTA-----PSVSKQENFR 1040
Qy 981 STCHTKERLPSHDFLAELKAPVNVGSIAVCORIQCDIPFFGIQEBFNATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMAEYFIVNTRVMNRT 1080
Qy 1041 YIKTSHNHLIVSFAELFNDVSFTLLPGQAFVRSOTETKVPFFPNPLPLIVSSVG 1100
Db 1081 FFASTFOTVQLTAAEIITNPOLFVIEENAVTIPLIMKPTKAEVPT--GVIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLALLAMTAGLWKLGFFKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 21-Jul-2000
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: GB:X17033; NID:g33906; PIDN:CRA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
Biochem. J. 279, 419-425, 1991
R;Cacimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIIb and GPIIc)
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <EXT>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 26.9%; Pred. No. 1.3e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDTENAMTFQ-ENARGFCQSVQL---QGSRVVVGAPQEIIVAAQNRGLYQC--DYST 54
Db 30 YNVGLPEAKIFSGPSFQGYAVQQFINPKGNLLVGSPPWSPGFPENRMGDVYKCPVDLST 89

```

```

Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYYKGLC 107
Db 90 ATCEKLNLTSTSPNVTETKNTMSLGLILTRNMGTFGLTCGPLWAQCGNQYTTTGC 149
Qy 108 FLFGSNLRQOPQKFPPEALRGCEQSDIAFLVDGSGSIIIPHPRRAKEPISVTWEOJK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDSENSIYPWD--AVKNFLEKFPVQGLDIG 205
Qy 166 KSKTIFSLMOYSEEFRIHFTFEFQNNPNRSLIKPITQLLG-RTHTATGIRKVVRELFN 224
Db 206 PKTVQGLLOYANNPRVFNLTNTYKKEEIVATSTQSYGGDLTNFPAIQAIVARKYVS 265
Qy 225 ITNGARKNAFKILITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFRSEK 279
Db 266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNDHNLIRFGIAGVLGNALDNTKN 324
Qy 280 SRQELNTVASKPRDHVFOINNFEALKTIONQLREKIPAIETGTOTQSSSSFEHMSQEGF 339
Db 325 LIKEIKAIASIPTEYFFNVSDAALLEKAGTLGEQIFSIETGVQG-GDNFQEMSQVGF 383
Qy 340 SAAITSNGP--LLSTVGSYDMAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGVAAL 394
Db 384 SADYSSONDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSTLGSVA 443
Qy 395 ILNRNVQSLVGLGAPRYQHIGLVAMFRQNTGMWESNANV-----KGTQIGAYEGASLC 448
Db 444 ISTGESTHEFVAGAPRANTGTQIVLYSVN-----ENGNIIVIAHQRDQIGSYFGSVL 498
Qy 449 DVDSNGSTDLVLGAPHYEQTR--GGQVSVCPPLRGQARWQCDVAVLGEQGPWGRFG 506
Db 499 DVDKDTITDVLVGLGAPMYMSDLKKEGRVYLFTIKKGIHQH--FLEGPEGIENTREG 555
Qy 507 AALTVLGVNGDKLTDVAIGAPCEENRGAIVLFGHTSGSGISPSHSORIASG--KLSPR 564
Db 556 SAIALSDINMGFNDVIVGSPLENQNSGAVTIYNGHQT-IRTKYSQILSGDGAFRSH 614
Qy 565 LQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQFVLVRKVAIMEFNEPREVARNYFEC 624
Db 615 LQYFGSLDGYDGLNGDSITDVSIGAFQVQVQLWSQIADVAIEASTPEKI--TLVNKN 672
Qy 625 DQVVKGEAGEVRVCLHVQKSTRDRRLREQIQSVVTYDIALD-----SGRPHSAVAVNETK 680
Db 673 AQII-----LKLCP---SAKFRPTKQNNQVAIVYNTILDADGFSRVTSGRLPKENN 721
Qy 681 NSTRRQTVLGLTQTC--ETLKLQPNCTEDPVSPIVLRNLSLVCTPLSAFGLNRPVLA 738
Db 722 ERCLQKNMVNQAQSCPEHIIYQEPS---DVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQLRFTALFPFKKNCNDNICQDDLSITF-----SFMSLDCLLVGGPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEDGLCISDLVDVRQIPAAQEQPFIVSNQKRLTFSVLKNK 832
Qy 795 GEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWRLACASAST-EVSGALKSTSCSINH 853
Db 833 RESAVNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAAASQKSVACDVGY 880
Qy 854 PIFPENSEYFNTTFDVSQKASLGNKLLKANVTSENMPRTNKTFFOLELPVKYAVYMW 913
Db 881 PALKREQQVTFINFDNLQ-NLQNAQSLSPQALSESQEBENKADNLVNLKIPILYDAEI- 938
Qy 914 VTSHGVSSTKYLNFNTASENSTRVQHQYQVSNLQOR-----SLPISLPLV----- 958
Db 939 ---HLTRSTNINFIYESSDGNVPSIVHSFEDVGPKEIFSLKVTGTSVPVSMATVHIPO 995
Qy 959 -----PVRINQTVIMDRPQVTF-SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGT---SSSVSPKSENFR---HTKE- 1047
Qy 989 LPSHSDFLAELKAPVNVCSIAVCQRIQCDIPFFGIQEBFNATLKNLSFDWYIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCLWKDVMHMKGYFVNVTTRINWNGTFASFTQT 1091

```


Db 101 RKNPNTCEQLQSGSPGSGCGTKCLEERDNQWLVTLRSQPGENGSIIVTCGHRWKNIFY 160
QY 96 TCSNTYVKGCLFLGSLNLRQPKQKPEALRGCPQEDSDIAFLVDGSGSIIPDFRRAKE 155
Db 161 MKSDNKLPTGICYWMPDLRLTSLK-----185
QY 156 FISTVMEQLKSKTLFSLMOYSBEFRIHFTFKFQNNPNRSLIKPITQLLGRTHATGI 215
Db 186 -----RMAPCYKDYT-----195
QY 216 RKVRELFNITNGARKNAFKLILITDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDADF 275
Db 196 -----RKFG-----200
QY 276 RSEKSRQELNTASKPPDRIVFOINNPEALKTIQNLREKIFAIEGTQTGSSSFEHMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGFSAAITNGPBLSTVSGYDNAGGVFLY--TSKEKSTFINMTRVDSMDNDAYLGYA-- 391
Db 206 QAGISSPYTQDLIVMGAPGSSYWTGTVFVYNNITNQYKAFVD--RQNVKFGSYLGSVG 263
QY 392 AAIIILNRVQSLVLGAPRYQHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCS 447
Db 264 AGHFRSPHTTEVVGGAPQHBOIGKAYIF-----SIDENELNVYEMKKGKLSYFGASVCA 319
QY 448 VDSDNSGTDLVLIGAPHYYEQTRGGQVSVCPLPRGORA--RMQCDVAVLYGEOQPGWRF 506
Db 320 VDLNADGFSDL-LVGPMPQSTIEBGRVFY--INSGMGAVMVNERVLVGSDDKYA-ARFG 376
QY 507 AALTUVLDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHQRIAGSKLSPRLQ 566
Db 377 ESTANLGDIDNDGFEDIAIGAPQEDDLRGAVIYNGRV-DGISTYSQRIEQGOISKSLR 435
QY 567 YFGOSLSGGODLTMDGLVDLTGCA--QGHVLLRSQPLRVKAIMFENPREVARNVPECN 624
Db 436 MFGOSISGQIDADNGYVDVAVAFOSDSAVLLTRPWIIVEASLS-HPESVNRTRKPDCT 494
QY 625 DQVVKGEAGEVRVCLHVQKSTRDLRLEGQIQSVVTVYDLALDSGR---PHSRVAVF--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGEVPGYVLFVNVSLDVHRKAESPSPRYFNSGT 547
QY 680 KNSTRQTQVLGTQTCETLKLQLPNCIEDPVPSPVLRNLFSL-----VGTPLSAFNGLR 734
Db 548 SDVITGSIKSVSSGKCRTHQAFMRKQVDRDILTPHVEATYHLGHVITKRNTEEPPLQ 607
QY 735 PVLAEDAQR-LFTALPFPEKNCNDNICQDDLSITFS-----FMSLDCLVGGCPREFNV 787
Db 608 PILQOKKEKDVIRKMINFARFCAYEN-CSADLQVSARVGLPKPYENKTYLAVGSMKTI 666
QY 788 TVTVRNDGDSYRTOVTFPPPLDLSYRKVSTLQNRQSRSMWLACASASTEVSGALKST 847
Db 667 NVSLFNAGDDAYETTLNVQLPTGLTYFIKILDSEK-----QINCE-----VTESSGIVK-L 716
QY 848 SCSSINHPFIPENSEVFNITFDVDSKASLGKLLKANVTSEN--NMPRTNKTEFOLELP 905
Db 717 ACSLGIYVDRLSRIDISFLDVSSLSRAHEDLSISVHASCENEGELDQVRNRTLTIP 776
QY 906 VKYAVVWVTSBGV--STKYLNTASENTSRVNQHQ-----YOVNLSGQKSLP-ISLVFL 957
Db 777 LRYEV--MLTVHGLVNPSTSVYSGSSENEPETCMAEKLNLTPHVTGISMVAPNSVKIM 834
QY 958 VPVRLNQTIVWDRP--QVTFSENLSSTCHTKE-----RLPSHSDFLAELR 1000
Db 835 VP---NSFUPQDKLVNLDVQTTTQOCHFKPHVYGRCTFAQQGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVWVNCSTIA--VCQRIQCIDIPFFGQEEFNATLKGNSLPDWYIKTSHNHLIVSTAAIL 1058
Db 892 KR-LLCYMKADQHCLOFLQN---FGKWSG-----KEASVHIQLEGPSIL 933
QY 1059 FNDOSVTLTLPQOGAFVRSQTEKVPFEPVNP-----1090
Db 934 EMDETSSL-----KFEIKATAPPEPHPKVIELNKDENVAVHVFLEGLHHORPKRHF 983

QY 1091 -LPLIVSVSGVGGILLIALITALYALYKLGFFKFRQVKMMSE 1128
Db 984 TIIITITISLLGLIVLLISCVWMAKAGFFKFRQVKSILQE 1022

RESULT 13

S06046

Integrin alpha-4 chain precursor - human

N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C:date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 24-Sep-1999
R:Accession: S06046; A39355; D28018
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989

A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other int
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046

A:Molecule type: mRNA

A:Residues: 1-1038 <TA>

A:Cross-references: GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1; PID:g33946
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655

A:Accession: A39355

A:Molecule type: DNA

A:Residues: 1-93 <ROS>

A:Cross-references: GB:M62841

R:Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A:Title: The very late antigen family of heterodimers is part of a superfamily of mole
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018

A:Molecule type: protein

A:Residues: 40-50, 'E', 52-53 <TA2>

C:Genetics:

A:Gene: GDB:ITGA4; CD49D

A:Cross-references: GDB:128032; OMIM:192975

A:Map position: 2q31-2q32

C:Superfamily: integrin alpha-4 chain

C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match

Best Local Similarity 10.8%; Score 635; DB 2; Length 1038;

Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAFRSEKSRQELNTVASKPPRD-----HVFOINNPEALKT-----ION 310
Db 121 GKTCLLEERDNQWLVTLRSQPGENGSIIVTCGHRWKNIFYIKENKLP TGCGYCVPPDLRT 180
QY 311 QUREKI-----FAIEGTQTGSSSFEHMSQEGFSAAITNGPBLSTVGSY 356
Db 181 ELSKRIAPCYQDVYKKGFGENFA-----SCQAGISSPYTKDLIVMGAPSS 225
QY 357 DNAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGYA---AAIILNRVQSLVGLAPRQ 411
Db 226 YWTGSLFVYNNITNKYKAFLDKQNVKF---GSYLGSYVGAGHFRSQHTTEVVGGAPQHE 282
QY 412 HGLVAMFRONTGMWESNANV---KGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYY 467
Db 283 QIGKAYIF-----SIDEKELNIHEMKKGLSGYFGASCAVDNADGFSDL-LVGPAMQS 337
QY 468 EOTRGQGVSVCPPLRGQRARWQC-DALVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIG 526
Db 338 TIREERGVFVY-INSNGSAGVAMNMTNLVGSDDKYA-ARFGESIVNLGIDINDGDFEDVAG 395
QY 527 ARGEEDNRGAVYLFHGTSGSGISPSHQRIAGSKLSPRLQYFGOSLSGGODLTMDGLVDL 586
Db 396 APQEDDLQAGAIYINGRA-DGISSTFSQRIEGLQISKLSLMSFGQSTSGQIDADNNGYVDV 454

A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: Z21035
A;Accession: T31437
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1041 <HER>
A;Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;
Best local similarity 24.8%; Pred. No. 3.9e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAITNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DMNDAYLG 389
DB 181 QAGFGIIFSDNSALVNGAFSGSYVLOQIYVQSLNRSV--VQATQESNTGTSYFDSNSYRG 239
QY 390 YAAAI--ILNRVQSLVLAGAPRYQHI--GLVAMFRQNTGMWESNANVKGTOIGAYFGASLC 446
DB 240 YSLALGDPNGGVQDYVVGTPRAESLMLVAIFDQNLNQFN---QVWGTOIVAYFGYSVT 296
QY 447 SVVDNSGSTDLVLIGAPHYYEQTGRGQSVCPPLPRQARWQCDVLYGEQ----- 498
DB 297 VVDI--NNDTYDILLVAGAPMYMDGPAIQ-----RWEGAVVYVYLQNPVGPQA 343
QY 499 -----GQWCRGGAALTVDGVNGDKLTDVAICAPGEEDNRGAVYLFHGTSGS 546
DB 344 SNRLSSTLIIGQIRSGRLSGIASIGDSNQDGFNDVAICAPYEGDDAGAVIYHG--SAN 402
QY 547 GISPSHSQRIAGSKLS--PRLOYFGQSLSGGQDITMDGLVLTVGAQ--GHVLLRSQPV 603
DB 403 GLKSTPAQVLTPTSLHSGITTFGFSLQGGQDMKKNYPDLLVGAESANTAVLIRTPVV 462
QY 604 RVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSV----- 658
DB 463 SLDATLNTPE-----IGINLENKTYE--LADGTMTVTSFIAMT 497
QY 659 -----VTYDLDLDSG--RPHSRVAFNETKNSTRQTVQLGL--TQTCETLKLQ 702
DB 498 CFTYTGNYLDPHDIDISYTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQFCPLRAY 557
QY 703 LPNCIEDPVPSPVILRLNFSLVGTPLSAFGN-----LRPVLAEADAQRLLTALPPEK 753
DB 558 VGNSEDKLTPIKVTLOYDL-----NDESRLOPHEILPIIDMATMSTQTKQVSIQN 609
QY 754 NCGNDNICQDLSITFSFMSLCLVGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSY 813
DB 610 NCVN--NICIPDLQVTVT--PNLPNIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLQLQF 667
QY 814 RKVSTLQNRQSRWRACESASSTVSGALKSTCSINHPFIPFN-----SEVTFNIT 867
DB 668 VRL-----ERKANMDFSTCSESD-----LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717
QY 868 FDVDS-----KASLGKLLKKNVTSNNMNPRTNKTEFQLELPKYAVYVMTVSHGVSTK 922
DB 718 GDXDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVCTCLLLSASYPEIVMYSTQED 775
QY 923 YL-----NFTASENTSRVMOHQVY-----SNLQORSIPIS-----LVFLVPV 960
DB 776 YVVPVPPFAKASEADIGMEVM--HLYEVRNTGSSNAGEVSLNIQWPKQKDEGEVLYLLGI 834
QY 961 RLNQTVIWRPQ-----VTFSENLSSTCHTKERLPKSHSDFLAELRKAPVV 1005
DB 835 MTEBEGVTCQLTQKANKPEGVKLEPSTKAKLSNSTTQVSGRRKREPEVAEALQTDN--VI 892
QY 1006 NCSIAVCQRIQCDDIPFGIOEEFNAT-----LKGNLSPDWYIKTSHNLLIIVSTAEILF 1059
DB 893 YCASDSCVLINCTI-----DEINASKSKVRIILGRF--W-----ERTF 928
QY 1060 NDSVFTLLPGGAGFVRQSTETKVE--PFEVFPNP-----LP----- 1092

Search completed: November 25, 2003, 14:22:04
Job time : 18.9579 secs

DB 929 QKAVSELTPVQVQATIASSASAAVKTIPTYNIPLPDROFSDSTKASTLVTTEELVPPVPTPIAW 988
QY 1093 -LIVGSSVGGLLLLALITAAALYKLGFFKR-----QYKQMMMS--EGGPP 1132
DB 989 WIIIVSVLGGIILLIILGLWKCGFFERKKPGEKEYAPVASADKGGPP 1038

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5846.5	99.5	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4470	76.1	1153	1 ITAM_MOUSE	P05555 mus musculus
3	3469	59.0	1163	1 ITAX_MOUSE	P20702 homo sapien
4	3411	58.0	1162	1 ITAD_HUMAN	Q13349 homo sapien
5	1557.5	26.5	1170	1 ITAL_HUMAN	P20701 homo sapien
6	1538.5	26.2	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1161.5	19.8	1167	1 ITAE_MOUSE	Q50677 mus musculus
8	1153	19.6	1179	1 ITAH_HUMAN	P38570 homo sapien
9	1103.5	18.8	1151	1 ITAI_HUMAN	P56199 homo sapien
10	1093.5	18.6	1189	1 ITAH_HUMAN	Q9UKX5 homo sapien
11	1085	18.5	1170	1 ITA2_BOVIN	P53710 bos taurus
12	1071	18.2	1178	1 ITA2_MOUSE	Q82459 mus musculus
13	1068	18.2	1181	1 ITA2_HUMAN	P17301 homo sapien
14	1064.5	18.1	1167	1 ITAG_HUMAN	O75578 homo sapien
15	1062	18.1	1180	1 ITAI_RAT	P18614 rattus norv
16	663	11.3	1039	1 ITA4_MOUSE	Q00651 mus musculus
17	635	10.8	1038	1 ITA4_HUMAN	P13612 homo sapien
18	630	10.7	1035	1 ITA9_HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA3_XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3_CRISP	P17852 cricetidae
21	567.5	9.7	1053	1 ITA3_MOUSE	Q62470 mus musculus
22	555.5	9.5	1053	1 ITA5_MOUSE	P11688 mus musculus
23	545.5	9.3	1050	1 ITA5_XENLA	Q06274 xenopus lae
24	542.5	9.2	1034	1 ITAV_CHICK	P26008 gallus gall
25	537.5	9.1	1130	1 ITA6_HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITA6_MOUSE	P43406 mus musculus
27	532.5	9.1	1066	1 ITA3_HUMAN	P26006 homo sapien
28	532	9.1	1049	1 ITA6_HUMAN	P08648 homo sapien
29	531.5	9.0	1072	1 ITA6_CHICK	P26007 gallus gall
30	526	9.0	1048	1 ITAV_HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8_CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6_MOUSE	Q61739 mus musculus
33	498	8.5	1179	1 ITA7_MOUSE	Q61738 mus musculus

ALIGNMENTS

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Hartsis P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Fahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	496	8.4	1396	1	ITA2_DROME	P12080 drosophila
35	491.5	8.4	1146	1	ITAI_DROME	Q24247 drosophila
36	491	8.4	1033	1	ITAB_MOUSE	Q9QM0 mus musculus
37	489.5	8.3	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.3	1025	1	ITAB_HUMAN	P53708 homo sapien
39	486	8.3	126	1	ITAM_CAVPO	P11578 cavia porce
40	476	8.1	1181	1	ITAT_HUMAN	Q13683 homo sapien
41	470	8.0	1106	1	ITAT_RAT	Q63258 rattus norv
42	467	7.9	1226	1	PAT2_CAEEL	P34446 caenorhabdi
43	445.5	7.6	1139	1	INAI_CAEEL	O03600 caenorhabdi
44	424	7.2	1115	1	ITAT_DROME	O44386 drosophila
45	385	6.6	1000	1	ITAT_DROME	O9WIM8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=5339202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
RT conservation across species and homology to platelet IIB/IIIA.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18)." ;
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Wildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment." ;
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface." ;
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- CELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFPA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide cdlib entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdlib.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S5227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAU51960.1; -
DR PIR; A31108; RWHU1B.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1M1U; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01813; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

Query Match		99.5%;	Score 5846.5;	DB 1;	Length 1152;		
Best Local Similarity		99.2%;	Pred. No. 0;				
Matches 1128;		Conservative	7;	Mismatches	1;	Indels	1; Gaps 1;
QY	1	FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI	60				
DB	17	FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI	76				
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGFLCFLGSLNLRQPOK	120				
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGFLCFLGSLNLRQPOK	136				
QY	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHPDFRRAKEFISTVMEQLKSKTFLSLMQYSEEF	180				
DB	137	FPFALRGCPQEDSDIAFLVDGSGSIIPHPDFRRAKEFISTVMEQLKSKTFLSLMQYSEEF	196				
QY	181	RIHFTPKFONNPNRSLKIPITQLLGRTHATGIRKVVRELNIINGARKNAFKILILI	240				
DB	197	RIHFTPKFONNPNRSLKIPITQLLGRTHATGIRKVVRELNIINGARKNAFKILIVI	256				
QY	241	TDGEKFGDPLGYEDVPEADREGVIRVVGVDGAPRSEKSRQELNTVASKPPRDHVFQIN	300				
DB	257	TDGEKFGDPLGYEDVPEADREGVIRVVGVDGAPRSEKSRQELNTVASKPPRDHVFQIN	316				
QY	301	NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG	360				
DB	317	NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG	376				
QY	361	GVFLYTSKEKSTFINTRVDSNDNDAYLVYAAAIILNRVQSVLVLGAPRYOHIGLVAMFR	420				
DB	377	GVFLYTSKEKSTFINTRVDSNDNDAYLVYAAAIILNRVQSVLVLGAPRYOHIGLVAMFR	436				
QY	421	QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQOVSCPL	480				
DB	437	QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQOVSCPL	496				
QY	481	PRQORARWQCDVLYGEGQPGWRCFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF	540				
DB	497	PRG-RARWQCDVLYGEGQPGWRCFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF	555				
QY	541	HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLRSQ	600				
DB	556	HGTSGSGISPSHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLRSQ	615				
QY	601	PVLVRKAIMEFNEPREVARNVFECNDQVVKGEAGEVRVCLVHVKQSTRDLREGQIOSVVT	660				
DB	616	PVLVRKAIMEFNEPREVARNVFECNDQVVKGEAGEVRVCLVHVKQSTRDLREGQIOSVVT	675				
QY	661	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	720				
DB	676	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	735				
QY	721	SLVGTPLSAFNLRLPVLAEADAQLFTLFPFKKNCNDNITCQDDLSITFSFMSLDCLVWG	780				
DB	736	SLVGTPLSAFNLRLPVLAEADAQLFTLFPFKKNCNDNITCQDDLSITFSFMSLDCLVWG	795				
QY	781	GPREFNVTVVRNDGEDSYQTQVTFPPDLSTRKYSTLQNRQSRWRACASASTEV	840				
DB	796	GPREFNVTVVRNDGEDSYQTQVTFPPDLSTRKYSTLQNRQSRWRACASASTEV	855				
QY	841	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF	900				
DB	856	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF	915				
QY	901	QLELPVKYAYVMVYTSHGVTSKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLPV	960				
DB	916	QLELPVKYAYVMVYTSHGVTSKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLPV	975				
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP	1020				
DB	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP	1035				

QY	1021	FFGIQEEFNATLKNLSFDWYIKTSHNHLIIYSTABILLPNDVSFTLLPQOGAFVRSQTET	1080
DB	1036	FFGIQEEFNATLKNLSFDWYIKTSHNHLIIYSTABILLPNDVSFTLLPQOGAFVRSQTET	1095
QY	1081	KVEPEFVNPLPLIVGSSVGGLLLALITAAALYKLGFFKROKMDMSEGGPPCAEPQ	1137
DB	1096	KVEPEFVNPLPLIVGSSVGGLLLALITAAALYKLGFFKROKMDMSEGGPPCAEPQ	1152
RESULT 2			
ITAM_MOUSE	STANDARD;	PRT;	1153 AA.
ID	ITAM_MOUSE		
AC	P05555;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1).		
GN	ITGAM.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88312584; PubMed=3044779;		
RA	Pytela R.;		
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";		
RL	EMBO J. 7:1371-1378(1988).		
RN	[2]		
RP	SEQUENCE OF 11-45 FROM N.A.		
RC	STRAIN=BALB/c; TISSUE=Spleen;		
RX	MEDLINE=86287312; PubMed=2942940;		
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,		
RA	Larson R.S., Roberts T.M., Springer T.A.;		
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).		
RN	[3]		
RP	SEQUENCE OF 17-28.		
RX	MEDLINE=85188276; PubMed=3887182;		
RA	Springer T.A., Teplow D.B., Dreyer W.J.;		
RT	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";		
RL	Nature 314:540-542(1985).		
CC	-!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLomerulonephritis. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.		
CC	-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.		
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.		
CC	-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.		
CC	-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.		
CC	-!- SIMILARITY: Contains 1 VWFA domain.		
CC	-!- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

-----
EMBL; X07640; CAA30479.1; -
EMBL; M14293; AAA39484.1; -
PIR; S00551; S00551.
HSSP; P11215; IABX.
MGD; MGI:96607; Itgam.
InterPro; IPR00413; Integrin_alpha.
InterPro; IPR02035; VWF_A.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00357; integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS02034; VWF_A; 1.
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Calcium; Repeat.
SIGNAL 1 16
CHAIN 17 1153
DOMAIN 17 1105
TRANSMEM 1106 1129
DOMAIN 1130 1153
REPEAT 31 84
REPEAT ? ?
DOMAIN 164 350
REPEAT 337 400
REPEAT 401 452
REPEAT 454 515
REPEAT 517 575
REPEAT 580 632
REPEAT 645 673
CA_BIND 529 537
CA_BIND 592 600
SITE 1132 1136
DISULFID 66 73
DISULFID 105 123
DISULFID 654 711
DISULFID 770 776
DISULFID 999 1023
DISULFID 1028 1033
CARBOHYD 58 58
CARBOHYD 86 86
CARBOHYD 391 391
CARBOHYD 696 696
CARBOHYD 734 734
CARBOHYD 772 772
CARBOHYD 801 801
CARBOHYD 881 881
CARBOHYD 907 907
CARBOHYD 941 941
CARBOHYD 980 980
CARBOHYD 994 994
CARBOHYD 1022 1022
CARBOHYD 1045 1045
CARBOHYD 1051 1051
CARBOHYD 1076 1076
SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;
Query Match 76.1%; Score 4470; DB 1; Length 1153;
Best Local Similarity 73.9%; Pred. No. 4.5e-295;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

```

QY 1 FNLDTENAMTFOENARGFQGSVQVQIGSRVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 60
 Db 17 FNLDTEHPMTFOENAKGFQGVVQVQIGTSVVVAAPQEAQVNTGALYQCDYSTSRCHPI 76

```

QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHOTCSNTYVKGCLFLFGLNLRQPOK 120
Db 77 PIQVPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVNGLCYLFGLNLRPPOK 136
QY 121 FPEALRGCPQESDIAFLVDGSGSIIPHDFRRAKEPISTVMSQOLKSKTFLPSLMQYSEF 180
Db 137 FPEALRGCPQESDIAFLVDGSGSIIPHDFRRAKEPISTVMSQOLKSKTFLPSLMQYSEF 196
QY 181 RIHFTPEFQNNPNPESLJKPITOLLGRTHRTATGRKVVRELFNITNGARKNAFKLIL 240
Db 197 RIHFTFNDPKRNPSPRSHVSPKQLNGRTKTASGRKVVRELFNITNGARKNAFKLIL 256
QY 241 TDGEKFGDPLGYEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHYFQIN 300
Db 257 TDGEKFGDPLGYEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHYFQIN 316
QY 301 NFEALNTIQNLRKEKIFAIEGTQSGSSSEFHEMSQEGFSAITSNGPLLSSTVGSVDWAG 360
Db 317 NFEALNTIQNLRKEKIFAIEGTQSGSSSEFHEMSQEGFSAITSNGPLLSSTVGSVDWAG 376
QY 361 GVPYLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNVOSILVLCAPRYOHIGLVAMPR 420
Db 377 GAFLYTSKOKVTPINTTRVDSMDNDAYLGYAAAILLRNVOSILVLCAPRYOHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSCPL 480
Db 437 ENGTWEPHTSINKSGIYGFASLCSVDMDADGNTNLILIGAPHYVEOTRGQVSCPL 496
QY 481 PROQRARWQCDVLYGEGQGPWGRFGAALTGLVDVNGDKLTDVAIGAPGEDNRGAYLFP 540
Db 497 PRG-RARWQCEALHGDQHPWGRFGAALTGLVDVNGDKLTDVAIGAPGEDNRGAYLFP 555
QY 541 HGTSGSGISPHSHSORIAGSKLSPLOYFGOSLGGDLTMDGLVDLTGAGQHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIGAHFSPGLQYFGOSLGGDLTMDGLVDLTGAGQHVLLRSQ 615
QY 601 PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLVHVKSTRDRREGQIQSVVT 660
Db 616 PVLRLTEATWEPSPKVARSVFACQEVKKNKDAAGEVRVCLVHVKSTRDRREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDVPSPVLRNLP 720
Db 676 YDLALDPVRSRIRAFDETNTNTRRRRTQVFLMKQKCTLLKLPDCVDSVSPILRLNY 735
QY 721 SLVGTPLSAGNLRPVLAEDAQRFLTALPFPEKNCNDNICQDDLSITFFSMSLDCLVVG 780
Db 736 TLVGEPLRSFGLNRPVLAMDAQRFFTFAMFPEKNCNDNICQDDLSITFFSMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWL-ACESASSTE 839
Db 796 GPQDFNMSVTLRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWL-ACESASSTE 855
QY 840 VSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTE 899
Db 856 GHGALKSTTNWNIHPANSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTE 915
QY 900 FQLELPKYAVYVWVTSYSHGVSTKYVLTASENTSRVWQHYQVSNLQORSPLISLVLP 959
Db 916 FQLELPKYAVYVWVTSYSHGVSTKYVLTASENTSRVWQHYQVSNLQORSPLISLVLP 975
QY 960 VRLNQTWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
Db 976 VQINNVTVWHPQVIFSONLSSACHTEQKSPHNFDFDQLETPVLNCSVAVCKRIQCDL 1035
QY 1020 PFFGIOEFNATLKNLSFDWYIKTSHNHLIIIVSTAILFNDSVFTLLPGQAFVRSOTE 1079
Db 1036 PSFNTQEIFNVTLLKGNLSFDWYIKTSHNHLIIIVSTAILFNDSVFTLLPGQAFVRSOTE 1095
QY 1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAAALKLGFKKQYKQDMSEGPPQAEQ 1137
Db 1096 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAAALKLGFKKQYKQDMSEGPPQAEQ 1153

```

[illegible]


```
QY 181 RIHFTKFEQNNPRLIKPIITOLLGRTHATGIRKVVRELFNTGARKNAFKILILI 240
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 198 QTHETFEFRRTSNPLSLLASVHOLQGFVTTATQNVHRLPHASTGARDATKILIVI 257
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 241 TDGKFGDPLGYEDVPEADREGVIRVIGVGDAPFRSEKSRQELNVTVASKPRPDHVFQIN 300
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 258 TDGKKEGSLDYKDVIPMAADAAGIIRYAIGVGLAFNRNSWKLNDIASQSEHIFKVE 317
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 301 NFEALKTIQNLREKIFAIBGTQTGSSSSFEHEMSQGFSAATISNCPPLISTVGSYDWAG 360
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 318 DFDALKIQNLKEKIFAIEGTETTSSSFELEMAQGFSAVFTPDGPFVLGAVGFTWSG 377
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 361 GVELYTSKEKSTENMTRVDSMDNDAYLGVAAAIIILNRVQSLVLGAPRYOHLGLVAMFR 420
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 378 GAFLYPNMSTFFINMQENVDMRDSYLGSTELALWKGVQSLVLGAPRYOHTGKAVIFT 437
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 421 QNTQWESNANVKGTQIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTGGQVSCPL 480
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 438 QVSRQWRMAEVTGTQGSYFGASLCSDVDNDSGSTDVLIGAPHYYEQTGGQVSCPL 497
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 481 PRGORARQCDAVLYGEGQPMGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAYILF 540
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 498 PRGWR-RWMCDAVLYGEGQPMGRFGAALTVLGVNGDKLTDVVVIGAPGEEENRGAYILF 556
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 541 HGTSGSGTSPSHSORIAGSKLSPLQYFGOSLSGGQDITWDLGLVLTGAGHVLILRSQ 600
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 557 HGVLGPSLSPSHSORIAGSKLSPLQYFGOSLSGGQDITWDLGLVLTGAGHVLILRSQ 616
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 601 PVLRVKAIMFENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 617 PVLVGVGSMQIPAEIPRSAFEQEVVSEQTQLVQSNICLVIDKRSKNLLGSLDQSSVT 676
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 661 YDLALDGRPHRAVNETKSTRQTVGLGTOTCETLKLQLENCEIDPVPVILRLNF 720
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 677 LDALDPRGLSPRATFQTKRSLRSRVVLGKACENFNLLPSCVEDSVTPITLRLNF 736
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALFPEKNCNDNICODDLSITFSPMSDCLVVG 780
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 737 TLVGKPLLAFLNLRPMLAALAQRYFTASLPPEKNCNGADHICODNLGISFSFPGUKSLVG 796
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 781 GPRFENVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 797 SNLELNAEVMWNGEDSYGTTIFSHPAGLSYRYVAEGQKQGLRSLHLTCDSPAVG-- 854
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 841 SGALKSTSCSINHPIFPENSVENTITFDVDSKASLGNKLLKANKVTSENMPRTNKTPEF 900
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 855 SGTWSTSCRINHLIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTRTSKTTF 914
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 901 QLELPVKYAVYVWVTSKYNLFTAS-ENTSRVMOHQOVSNLQORSIPISLVFLVP 959
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 915 QLELPVKYAVYVWVTSKYNLFTAS-ENTSRVMOHQOVSNLQORSIPISLVFLVP 974
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 960 VRLNQTIWDRPQVTFNSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQDI 1019
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 975 VELNQEAVMDEVSHQPNPSLRSSEKIAFPASDFLAHQKPNVLDCSIAGCLFRCDV 1034
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1020 PFFQTEBFNATLKNLGFQWYIKTSHNHLIVSTAEILFNDVSTLPPGQAFVRSOTE 1079
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1035 PSFSVQEBELDTLKNLGFQWYIKTSHNHLIVSTAEILFNDVSTLPPGQAFVRSOTE 1094
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1080 TKVEFEFVNPPLPLVGVSVGLLLALLIATAALYKLGFFKQYKDMSEE 1128
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1095 TVLEKYKVHNPTPLVGVSVGLLLALLIATAALYKLGFFKQYKDMSEE 1143
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 4
ITAD_HUMAN
ID ITAD_HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE GN Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
OS ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690(1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sp1 and Sp3.";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human
RT beta 2-integrin alpha subunit.";
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "alpha2 integrin is expressed on human eosinophils and functions
RT as an alternative ligand for vascular cell adhesion molecule 1
RT (VCAM-1)."
RL J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT binding interface between I domain and VCAM-1."
RL J. Immunol. 163:1984-1990(1999).
CC -1- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES
CC FROM THE BLOOD.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U37028; AAB38547.1; --
```

DR EMBL; U40274; AAB60634.1; --
 DR EMBL; U40275; AAB60635.1; --
 DR EMBL; U40276; AAB60636.1; --
 DR EMBL; U40277; AAB60637.1; --
 DR EMBL; U40278; AAB60638.1; --
 DR EMBL; U40279; AAB60639.1; --
 DR EMBL; AF187881; AAF62875.1; --
 DR HSP; P11215; IABX.
 DR Genew; HGNC:6146; ITGAD.
 DR MIM; 602453; --
 DR GO; 0008305; C: integrin complex; TAS.
 DR GO; 0004895; F: cell adhesion receptor activity; TAS.
 DR GO; 0016337; P: cell-cell adhesion; NAS.
 DR GO; 0007160; P: cell-matrix adhesion; NAS.
 DR GO; 0006955; P: immune response; NAS.
 DR InterPro; IPR00413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00327; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS02234; VWF_A; 1.
 DR InterPro; IPR00413; Integrin_alpha.
 KW Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium;
 KW Magnesium.
 FT SIGNAL 1 17
 FT CHAIN 18 1162
 FT DOMAIN 18 1100
 FT TRANSMEM 1101 1124
 FT DOMAIN 1125 1162
 FT REPEAT 32 85
 FT REPEAT ?
 FT DOMAIN 150 332
 FT REPEAT 350 400
 FT REPEAT 401 452
 FT REPEAT 454 516
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 465 473
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1127 1131
 FT DISULFID 67 74
 FT DISULFID 106 124
 FT DISULFID 655 710
 FT DISULFID 769 775
 FT DISULFID 846 861
 FT DISULFID 994 1018
 FT DISULFID 1023 1028
 FT CARBOHYD 59 59
 FT CARBOHYD 87 87
 FT CARBOHYD 99 99
 FT CARBOHYD 391 391
 FT CARBOHYD 691 691
 FT CARBOHYD 733 733
 FT CARBOHYD 873 873
 FT CARBOHYD 957 957
 FT CARBOHYD 1046 1046
 FT CONFLICT 500 500
 FT CONFLICT 515 518
 FT CONFLICT 825 825
 FT CONFLICT 984 984
 FT SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;

Query Match 58.0%; Score 3411; DB 1; Length 1162;
 Best Local Similarity 59.4%; Pred. No. 3.3e-223;
 Matches : 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNRGSLYQCDSYTGSCPEI 60

18 FNLDVEEPTIFQEDAGGFGQSVVQLOGSRVVGAPLEVVAAANQTGLYDCAATGMCQPI 77
 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLPGSNLRQPOPK 120
 78 PLHIRPEAVNMSLGLSLAATTNGSRLLACGPTLHRVCGENSYSKSGCLLGGRW-EIQT 136
 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKKSKTLFSLMQYSSEF 180
 137 VPDATPECPHQEMDIVFLIDGSGSIDQDNDFNMKGQVQAVMGQFEGTDTLFFALMQYSNLL 196
 181 RIHFTFKEPQNNPNRSLIKPIITOLLGRTHATGTRKVVRELFINITNGARKNAFKILLI 240
 197 KIHFTFTQRTSPSQSLVDPIVLQKGLTFTATGILTVVTQTLPFHKNKGARKSAKILVI 256
 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDFRSEKSRQELNIVASKPRDHDVQIN 300
 257 TDGQKYKDPLEYSVDIPQAEKAGIIRYALGVGHAFCGPTARQELNITISSAPQDHFVKVD 316
 301 NFEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAAITSNGLPILLTVGSYDWAG 360
 317 NFAALGSIQQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 376
 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 377 GAFLYPPNMSPTFINNSQENVMDRDSYLGSTELALWKGQVNLVGLGAPRYQHTKAVIET 436
 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYVEOTRGGQSVCP 480
 437 QVSRQWRKKAETVGTQIGSYFGASLCSDVDVDSGSTDLLIGAPHYVEOTRGGQSVCP 496
 481 PRGQARMQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
 497 PRGQVQWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLSGGGDLTMGDLVDTVGAQGHVLLRSQ 600
 557 HGASESGISPSHSQRIAGSKLSPRLQYFCQSLSGGGDLTMGDLVDTVGAQGHVLLRSQ 616
 601 PVLRYKATIMEFNPREVARNVFECDQVVKGEVUCLHVQKSDRLRREGQISVVT 660
 617 PVLKGVAMRFPVEVAKAVYRCWEEKPSALEAGDATVCLTQKSLDQL--GDIQSSVR 674
 661 YDLALDSGRPHSRVFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 675 FDLALDPGLTSTRATFNETKNTLTERTLGLGHCETLKLLJLPCDVVDVSPILHLNF 734
 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFKNCGNDNICDDLSITFSWSDCLVVG 780
 735 SLVREPILSPQNLRLPVLAVGSDQLFTASLPPEKNCQDGLCEGLGVTLSFSLQTLTVG 794
 781 GPREENVTYVENDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 795 SSELNVIYVWNAGEDSYGVTVSYLYPAGLSHRRVSGAQKQPHOSALRLACETV-PTED 853
 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKYTSNNMPTWKTFP 900
 854 EG-LSSRCSNVNPIFHEGSGNCTFIVTDPVSKATLGLDRMLMRASSENKASSKATF 912
 901 QLELPVKYAVYVMTVSHGVSTKYLNF-TASENSTRVMOHOYQVNSLQGRSLPISLPLVP 959
 913 QLELPVKYAVYVMTVSHGVSTKYLNF-TASENSTRVMOHOYQVNSLQGRSLPISLPLVP 972
 960 VRLNQTIVDRQVTFSENLSSTCHTKERLPHSDFLAELRAKVPVNCISAVCORIQCDI 1019
 973 VLLNGVAVVWVMEAPSQSL--PCVSRKPPQHSDFLTQISRSPMLDSCSIADCLQPCDV 1030
 1020 PFFGQIEEFNATLKGNLSPDWIKYKSHNHLIVSTAEIILFNDSVFTLLPGQCAFVRSQTE 1079
 1031 PFSVQEELEDFTLKNLSFGWVRETLOKKVLVSVVAEITFTDSVSVQLPQCAFMAQME 1090
 1080 TKVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKRYQKMMSE 1128

Db 1091 MVLBEDEYNAIPIMSSVGALELLALITATLYKLGFKRHYKEMLED 1139

RESULT 5

ITAL HUMAN STANDARD: PRT; 1170 AA.

AC P20701; O43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).

DE ITGAL OR CD11A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=3537122;

RA Larson R.S., Corbi A.L., Herman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";

RL J. Cell Biol. 108:703-712 (1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=99425270; PubMed=10493829;

RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";

RL Genomics 60:295-308 (1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Qu A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281 (1995).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398682; PubMed=8805579;

RA Qu A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";

RL Structure 4:931-942 (1996).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=99425288; PubMed=10493852;

RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon llovasstatin binding to the CD11a I-domain.";

RL J. Mol. Biol. 292:1-9 (1999).

CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2, ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES AND MONOCYTES.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS: Event=Alternative splicing; Named isoforms=2; Name=1; IsoId=P20701-1; Sequence=Displayed; Name=2; IsoId=P20701-2; Sequence=VSP 002738; Note=No experimental confirmation available;

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL; Y00796; CAA68747.1; -.

DR EMBL; AC002310; AAC31672.1; -.

DR PIR; S03308; S03308.

DR PDB; 1LFA; 29-JAN-96.

DR PDB; 1ZON; 07-DEC-96.

DR PDB; 1ZOO; 07-DEC-96.

DR PDB; 1ZOP; 07-DEC-96.

DR PDB; 1CQP; 07-AUG-00.

DR PDB; 1DGO; 03-FEB-00.

DR PDB; 1MJN; 28-JAN-03.

DR PDB; 1MQ8; 14-JAN-03.

DR PDB; 1MQ9; 14-JAN-03.

DR PDB; 1MQA; 14-JAN-03.

DR Genew; HGNC:6148; ITGAL.

DR MIM; 153370; -.

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0006928; P:cell motility; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS00234; VWFA; 1.

DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; 3D-structure; Magnesium; Calcium; Repeat; Alternative splicing.

FT SIGNAL 1 25

FT CHAIN 26 1170 INTEGRIN ALPHA-L.

FT DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1089 1112 POTENTIAL.

FT DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

FT REPEAT 92 91 FG-GAP 1.

FT REPEAT 42 149 FG-GAP 2.

FT DOMAIN 170 349 VWFA.

FT REPEAT 2 2

FT REPEAT 401 455 FG-GAP 3.

FT REPEAT 457 516 FG-GAP 4.

FT REPEAT 518 575 FG-GAP 5.

FT REPEAT 578 630 FG-GAP 6.

FT CA_BIND 468 476 FG-GAP 7.

FT CA_BIND 530 538 POTENTIAL.

FT CA_BIND 590 598 POTENTIAL.

FT SITE 1115 1119 GFFKR MOTIF.

FT DISULFID 73 80 BY SIMILARITY.

FT DISULFID 111 129 BY SIMILARITY.

FT DISULFID 653 707 BY SIMILARITY.

FT DISULFID 771 777 BY SIMILARITY.

FT DISULFID 845 861 BY SIMILARITY.

FT DISULFID 998 1013 BY SIMILARITY.

FT DISULFID 1021 1052 BY SIMILARITY.

FT CARBOHYD 65 N-LINKED (GLCNAC...) (POTENTIAL).

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1
alpha-subunit and its expression in COS cells.";
RL J. Immunol. 147:369-374(1991).
RN [2]
RP SEQUENCE OF 24-42.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
glycoproteins and unexpressed relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
LEUKOCYTES RECRUITMENT.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VMFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC
CC EMBL; M60778; AAA39426.1;
DR PIR; I56126; I56126.
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VFWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium;
KW Repeat.
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT REPEAT 628 647
FT CA_BIND 466 474
FT CA_BIND 528 536

FT CA_BIND 588 596
FT SITE 1111 1115
FT DISULFID 70 77
FT DISULFID 108 126
FT DISULFID 147 199
FT DISULFID 651 705
FT DISULFID 767 773
FT DISULFID 840 856
FT DISULFID 993 1009
FT DISULFID 1017 1048
FT CARBOHYD 86 86
FT CARBOHYD 185 185
FT CARBOHYD 270 270
FT CARBOHYD 444 444
FT CARBOHYD 668 668
FT CARBOHYD 696 696
FT CARBOHYD 724 724
FT CARBOHYD 728 728
FT CARBOHYD 776 776
FT CARBOHYD 857 857
FT CARBOHYD 880 880
FT CARBOHYD 890 890
FT CARBOHYD 899 899
FT CARBOHYD 927 927
FT CARBOHYD 1056 1056
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489E8232F CRC64;
Query Match 26.28; Score 1538.5; DB 1; Length 1163;
Best Local Similarity 34.28; Pred. No. 3.6e-96;
Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;
Qy 1 FNLDTENAMTFOENA-RGFGQSVVQLQSGSRVVGAPQEIIVAANQRGSLYQCDYSTGCEP 59
Db 24 YNLDTRTQSLAQAGRHFGYQVLQIEDG-VVVGAPCE---GNDTGLYHCRTSSEFCQP 79
Qy 60 IRLQVPVEAVNMSGSLAATSPQLLACGPTVHQTCSNTYVYVKGCLFLFGSNLRQPO 119
Db 80 VSLH-GSNHTSKYLGMTLATDAKAGSLACDPLGSLRTCDQNTYLSGLCYLFPQSLGPM 138
Qy 120 KFPALRGCCQEDSDIAFLVDGSGSIIPHDFRAKEPISTVMQOLKSKTLFSLMOYSEE 179
Db 139 QNRPAYQECMKRGKVDLFLFDGSGSLDRKDFEKLFEKMDVMKLSLNTSYQFAVQFSTD 198
Qy 180 FRIHFTKEF-QNNPNPRSLIKPITQLLGRTHATGIRKVVRELFNTNGARKNAFKILI 238
Db 199 CRTEFTLDYVYKQKNDVLLGSGVQPMFLTNTRFVAVVYVHVFKEESGARDPATKVLV 258
Qy 239 LIIDGKEFGDPLGYEDVIPLEADREG-----VIRYVIGVDAPRSKSKOELMTVASKP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIGICKHFVSQKQKTLHFASEP 304
Qy 292 PRDHVFOINNPEALKTIQNLREKI FAIEGTQSGSSSEHEMSQEGFSAITSNGPLLS 351
Db 305 VEEFVKILDTFEKLDLFTDLQRIYAIEGTNRDLTSFNNELSSSGISADLSKGHAVVG 364
Qy 352 TVGSDYDWAGGVF-LYTSKEKSTFTINMTRVDSMDMDAYLGAAAA-IILNRNVQSLVLGAPR 409
Db 365 AVGAKDWAGGFLDLREDLQGATFVQEPFLTSVGRGGLYGYTVAMWTSRSSRPLLAAGAPR 424
Qy 410 YQHIGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYY 467
Db 425 YQHVQVLLFQAPGAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGAELLLIGAPLFF 484
Qy 468 EOTRGQSVSCPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGVNVDKLTDAIGA 527
Db 485 GEQGRGVFTY---QRRQSUFEMVSELQGDGPGYPLGRFGAAITALTDLINGRLTDAVGA 541
Qy 528 PGEEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSRPLQYFGQSLSGQDLTDLGLVDLT 587
Db 542 PLEE--QGVAVIFNGKPG-GLSPQSPQRIQAGVFPGRVWFGRIHGVKDLGGDRLDV 598
Qy 588 VQAQHVLLLSQVPLRVKAI MEFPNPREVARNFECDNQVYVKGKAG-EVRVCLVHKVST 646

FT CARBOHYD 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; 88331C115DCCCFD CRC64;

Query Match 19.8%; Score 1161.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 1.4e-70;
Matches 357; Conservative 214; Mismatches 457; Indels 205; Gaps 43;

QY 1 ENLDTENA--MTFQENARGFGQSVVQLOGSRVVVGAQEIIVAAANORG--LYOCDY 52
DB 20 FNMVDWAWTALQPGAPAVLSSLLHLDPSN-----NOTCLLVARRSSNRNTAALYRCAL 74
QY 53 STGSCERIRLOPVVEAVNMSLGLSLAATT--SPQLLAC-GPTVHQCSTENTYVKGCLFL 109
DB 75 SI-SPDEIACO-PVEHICMPKRGYQGVTLVGNHNGVLVCIQVQARKERSLNSLTGACSL 132
QY 110 FGSNLRQOPQKPEALRG-----C-----PQE 131
DB 133 LTPNLDLQAQVYFSDLEGFLDPGAHVDSGDYCRSGKSGSTGEKKSARRRRTVBEDEBED 192
QY 132 DSDIAFLVDGSGSIIPHDFRAKPISTVMEQL--KSKTLFSLMOYSEEPRIHFTPEF 189
DB 193 GTEIAFLVDGSGSGFDFQAKNPSTMMNFVEKCECPNFPALVQYGAIVTQEDFQES 252
QY 190 QNNPNRSLIPIITQLGRTHATGIRKVVRELFINITGARKNAFKILILITDGEKFGDP 249
DB 253 RDINASLAKVOSIVOVKEVTKTASAMQHVLDNIPTPSGRKAKLVWVLTDGIDFGDP 312
QY 250 LGYEDVPEADREGVIRVIVGVDAFRSEKQELNTVASKPRPDHVFQINNFALKTIQ 309
DB 313 LNLTVINSKMGVVRFAIGVDRFKNNTVRELKLIASDPKEAHTKVTNYSALDGLL 372
QY 310 NQLRKIEPIAETQGTGSSSEHEHMSQGSNAITSNGP-LLSTVGSVDWAGGVFLY-TS 367
DB 373 SKLOORIYHMEGT--VGDALQYLAQTGFSAQILDKGOVLGTVGAFNWSGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DNMDAYLGAAYAILLRNRVQSLVGLAPRQVHIGLVAMFRONTGM 425
DB 430 NCRGFLNQAKEDRTVQYSLGSLAVLHKAGISVAGNPRHLKAGVFLKEDR- 488
QY 426 WESNA---NVKQTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYVEQTRGQGVSCPLPR 482
DB 489 -BEDAFVRIEIQMGSGVFGSLCPVDIDMDGTTDFLLVAAPFYHIRGEGRVYVQYPE 547
QY 483 GQARWOCDAVLGYGQGPWGFGAALTIVLDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGLTNSRFGFAMAAGVINDQKFTDVAIGAPLEGFGAGDGASYS 606
QY 537 VYLFHGTSGSGISPSHSORISAGSLSPRIQYFGSLGGQDLTMDGLVLTGAGHVL 596
DB 607 VYIYNGHSG-GLYDSPSQIRASSVASGLHYFGMSVSGGLDPNGDGLADITVGSRSADV 665
QY 597 LRSQVLRVKAIMENPREVARNVFECDNQVVKGEAGEVRVCLHVQKS---TRDRLEG 653
DB 666 LRSRPVDLTVSMTEFP-----DALPMVFIRGM--DVNLCFEVDSSVVAEPGLREM 715
QY 654 QIQSVVYTDLALDSGRPSRVNETKNSRTOQVILGLTQFC----- 696
DB 716 FLNFTVDV-----TKQRQLQCDSDSGCQCLRWNGSGFLCBHFWLI 760
QY 697 ETLLQLQPCIEDPVPVILRNLFSLVGTPLSAFNGLR-----PVLAEADQRULTALF--P 750
DB 761 STEEL-----CEDCFSNITIKTYE-----FOTSGRRDYPNPTL--DHYKEPSAIFQLP 809
QY 751 FEKNGNDNICODDLSITFSFMSLDCVLVVGGRPRENVTVTVNNGEDSVRTQVTFPPFLD 810
DB 810 YEKQCKGNKVFCAEIQIOLATN-ISQBELVVGVTKEVTNLSITNSGDSYMTNMLNYPN 868
QY 811 LSVRKVSTLQNRQSRWPLACESASSSTVSALKSTCSINHPFIPENSEVTFNITFDV 870
DB 869 LQFKKI-----QKPSRPVQCDPKPV---ASVLVNMCKIGHPII-LKSSVNVSVTWQL 918
QY 871 DSKASLGNKLLKANVTSENNPRTNKTEFQLELPVKYAVYVWVTVSHGVSTKYLNTASE 930

DB 919 EESVFNRTADITVTISNSNEKSLARETR---SLOPRHAFIAVLSR--PSVMTMN--TSQ 971
QY 931 NTSRVWQHQQVSNLQORSIPLSLVFLVPLVRLNQTVIMDRPOVTPSPENLSST-----CHT 985
DB 972 SPDSHKKEFFNVHGENLFGAVFQLOICVPFKQDF-----QIVRVNLTKTQDHTECTQ 1025
QY 986 KERLPSHSDFLAEIRKAPVNVCSIAVCORIQCDIPFGIQEEFNATLKGNSLFDWYIKTS 1045
DB 1026 SQEPACGSDPVQVHKWHSVVCAL-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILFNDVSFTLLPQCGAFVRSQETKVEPF-----EVPNPLPLIV 1095
DB 1066 HTRQLLRDVSLEPILGEISFNKSLYBGLNAE-----NHRKITVIFLKEEETRSPLII 1119
QY 1096 GSSVGGLLLALITAAALYKLGFFKROVKMWS 1128
DB 1120 GSSIGGLLVLLVIAILFKGFFRKIQQLNLE 1152

RESULT 8
ITAE HUMAN
ID ITAE_HUMAN STANDARD; PKT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RX MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Aniketer Y., Dietrich N.L., Maduro V.V., McDowell G., Shotelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
CC EPITHELIAL CELLS.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25851; AAB59359.2; -
CC EMBL; AF168787; AAF43107.1; -
CC PIR; A53213; A53213.
CC HSSP; P11215; LA8X.
CC Genew; HGNC:6147; ITGAE.
CC MIM; 604682; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 3.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; WFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
CC SIGNAL 1 18
FT CHAIN 19 1179 INTEGRIN ALPHA-E.
FT CHAIN 19 177 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1125 1147 POTENTIAL.
FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 181 198 GLU-RICH (ACIDIC).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT 145 199 X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 200 391 VWFA.
FT REPEAT 401 456 FG-GAP 3.
FT REPEAT 457 506 FG-GAP 4.
FT REPEAT 510 571 FG-GAP 5.
FT REPEAT 573 638 FG-GAP 6.
FT REPEAT 641 693 FG-GAP 7.
FT CA_BIND 522 530 POTENTIAL.
FT CA_BIND 586 594 POTENTIAL.
FT SITE 654 662 POTENTIAL.
FT SITE 1150 1154 GFFR MOTIF.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 126 159 BY SIMILARITY.
FT DISULFID 706 762 BY SIMILARITY.
FT DISULFID 823 829 BY SIMILARITY.
FT DISULFID 893 907 BY SIMILARITY.
FT DISULFID 1008 1033 BY SIMILARITY.
FT DISULFID 1041 1057 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	934	934	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	954	954	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1065	1065	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	360	360	D -> E.
FT	VARIANT	1041	1041	/FTID=VAR_008884.
FT	MUTAGEN	208	208	/FTID=VAR_008885.
FT	MUTAGEN	316	316	D->A; LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	477	477	E->A; LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	482	482	V -> I (IN REF. 3).
FT	CONFLICT	950	950	Q -> R (IN REF. 3).
FT	CONFLICT	1019	1019	R -> W (IN REF. 3).
FT	CONFLICT	1019	1019	A -> V (IN REF. 3).
SQ	SEQUENCE	1179	1179	AA; 130088 MW; E558902EDP9D95E1 CRC64;

Query Match 19.6%; Score 1153; DB 1; Length 1179;
Best Local Similarity 29.1%; Pred. No. 5.3e-70;
Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;

QY	45	GSLYOCYSTGS--CEPI-RLQVP-----VEAVNMSGLSLAATTSPPOLLACGPTVHQ	95
DB	65	GPLHRCSLVQDBEILCHPVHVPPIKGRHGRVTVVRSHHGVLICI-----QVLVRRP--HS	117
QY	96	TCSENTYVKGCLFGLSGLNLROQPQ-----	119
DB	118	LSSELT--GTCSLGLGPDRLRPOAQNFFDLENLLDPDARVDGDCYSNKEGGEDDVNTA	174
QY	120	KFEALRGCPQED-----SDIAFLVDGSGSIIPHDPRRAKEFIITVMEQL--	164
DB	175	RORALEKEEEDDEEEDDEEAGTEIATILDGSGSIDPPDFORAKDFISNMNRNFEY	234
QY	165	KSKTFLSLMOYSEEFRIHFTKFNPNPRLIKPTQLLGRTHATGIRKVVRELN	224
DB	235	KCFECNFALVGVGVIGVITQTEFLRDSQDVNASLARVQNTQVGSVTKTASAMOHVLSIF	294
QY	225	ITNGARKNAFKILITDGEKFGDPLGYEDVPIPEADREGVIRYVIGCVGDAFSEKSRQEL	284
DB	295	SSHGSRKASKVMVVLTDGGIFEDPLNLTTVINSPKMQGVVERFALCVGEFSAKARTREL	354
QY	285	NTVASKPRDRHVQINFEALKTIQNLREKIFAEQTGTGSSSSFEHMSQEGFSAAT	344
DB	355	NLIASDPDETHAFKVTNYMALDGLLSKLRYNIISMEGT---VGDALHYQLAQIGFSAQIL	411
QY	345	SNGP--LLSTVGSYDNAGGVFLY--TSKEXSTFINMTFVDSMDNA---YLGVAAILLN	398
DB	412	DERQVLLGAVGAFDWSGGALLYDTRRRGRFLNQTAATAAADAEEAQAQSYLGVAVALHKT	471
QY	399	RVQSLVLAGAPRYOHIGLVAMFR--QNTGMWESNANY--KGTQICAYFGASLCSVDVDSNGST	456
DB	472	CSLSYVAGAPQYKHG--AVFELQKREASFLPVLEGEQMSYFGSELCPVIDIDMDGST	529
QY	457	DLVLIGAPHYBQTRGGQVSCPLPRGQARQWCDAVLVGEQGPWGRFGAALTVLGDVN	516
DB	530	DFLLVAAPFYHVHGEGRVYVYRLSE--QDGSFLARILSLHCHGFTNAREGFAMAAGDLS	588
QY	517	GDKLTDVAIGAP----GEEDNR--CAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQ	570
DB	589	QDKLTDVAIGAPLEGFGADDDGASFGSVIYNG--HWDGLSASPSQIRASTVAPGLQYFGM	647
QY	571	SLSGGQDLTMDGLVDLTVGAQGHVLLLRLOPVLKAIWFEFNPVARNVFECDQVVKG	630
DB	648	SNAGFDISGDLADITVGTLCQAVFVRSPVVRKLVSMAPFTSALP-----IGF	697
QY	631	KEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDALDLSGRPHSRVAFNETKNSRTRQVL	690
DB	698	NGVWVRLCFEI--SSVTTASESGLEALLNFITLDVDVGKRRRLQCSRVSRCLGUREWS	756
QY	691	GLTQTCETLKLQLPN-----CIEDPVSPVLRNFSLVGTPLSAFGLNLRPVLEADAORLFT	746
DB	757	SGSQLCEDL--LLMPTEGELCEDCFSNASKVKYSYL--QTPEGQTOHPQILDRYTPFAI	814

QY 747 ALFPKCKGNDNTCDDLSITFSFMSLDCLVWGCPREFNVTVTVRNDGEDSYRTQVTFP 806
DB 815 FQPYEKACKNKKFCVAELQLA-TTVSQQLWGLTKELTUNLNTSGEDSYMTSMALN 873
QY 807 PPLDLSYKRVSTLQNSQSRQSRWRLACSSASTEVSGALKSTSCSINHPPIFENSEVTFNI 866
DB 874 YPRNLQ-----LRMKQPPSPNIQCDDPQV---ASVLNMCNIGHPVL-KRSAAHSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN---MPRTNKTBFQ---LELPVKYAVYVTVTSHGV 919
DB 924 VQOLEENAFPNRTADITVTVNSNERRSLANETHTLQPHGFVAVLSPKSPMYNTGQGL 983
QY 920 S--TKYLNFTASENTRVQHQVQSNLQGRSLPISLVFLVPLVNLQTVIWDNRQVTFSE 977
DB 984 SHHKEFLPHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLLTRTQ 1028
QY 978 NLSSTCHTKPLGSHSDPLAELRKAPVNVVNCIAVCQRIQCIPFFGQIEEFNATLKGMLS 1037
DB 1029 ASTVCTWSQERACAYSS-VQVHEWHSVCVIA-----SDKENVTVAEIS 1073
QY 1038 FDMYIKTSHNLLIVST-----AEILFNDVSFTLLPQGGAFVRSQTKETKVEPEVNPL 1091
DB 1074 WD-----HSELLKDVTELQILGEISFNKSLYEGLNAENH--RTKITVWELKDEKVHSL 1125
QY 1092 PLIVGSSVGGILLILLALITAAALYKLGFFKRYQKDMWSE 1128
DB 1126 PIITKGSVGGLLVLIVLVILVFKCGFFKRYQQLNLE 1162

RESULT 9

ITAL HUMAN
ID ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit";
RL J. Biol. Chem. 268:2989-2996 (1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
CC PIR; A45226; A45226.
DR PDB; 1OC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGAI.
DR MIM; 192968; .
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.

DR PFam; PF01839; FG-GAP; 3.
DR PFam; PF00357; Integrin_A; 1.
DR PFam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
KW DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 FG-GAP 1.
FT REPEAT 16 75 FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA BIND 470 478 POTENTIAL.
FT CA BIND 552 560 POTENTIAL.
FT CA BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GPFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;

Query Match

Best Local Similarity 18.8%; Score 1103.5; DB 1; Length 1151;
Matches 344; Conservative 212; Mismatches 486; Indels 195; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQBIIVANORGSLYQCDYSTGS 56
DB 1 FNVDVKNMTSGPVEDMFGYTVQYENEKGVLLIGSLPVGQPKNRTGVDVYKCPVGRGE 60
QY 57 CEP-IRLQVPEA-----VNMSLGLSAAATSPPLQACGTVTQTCSENTYVYVGL 106
DB 61 SLPCVKLDLPVNTSIPNVTENKNTFGSTL-VTNPNPGFLACGPLYVRCGHLHYTTGI 119

QY 107 CFLFGSNLRQOPKFFBRLRGCPQEDSDIAFLVDGSGSIIPHDPRAKBFIITVMBQLK- 165
DB 120 CSDVSPFTQVNSIAP--VOECSTQ-LDIVILVDGSGNSIYPWDSVTA--FLNDLLKRMID 174
QY 166 -KSKTLFSLMOYSEEFRIHFTKFEFONNPNRSLIKIPITQLLGR-THTATGIRKVVRELF 223
DB 175 GPKQTQVIGVQGENVTHEFNLNKYSTSEVLVAKKIVQRGRTMTALGTDGTARKEAF 234
QY 224 NITNGARKNAFKILITDGEKFGDPLGYEDVIPEADREGVIYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKKVMVITDDESH-DNHLKVKVQDCEDENIQRFSAIILGYSYNGNLSTE 293
QY 279 KSRQELNTVASKPRDHVFOINNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQBG 338
DB 294 KFVEEIKTASEPTEHFFNVSDALVITVTLGERIFALEATADQSAASEMEMSQTG 353
QY 339 FSAAITNSGPLSTGVSXDWAGGVLYTSKE-----KSTP-INMTVDSMDNDVLYGAA 392
DB 354 FSAHYSQDWMLGAVGAYDWNGTVVMOKASIIIPRNTTFNVSTKKNPL-ASYLGTVV 412
QY 393 AITLNRVOSL-VLGAPRYQHIGLVAMFRONTGMESNANVKGTQIGAYFGASLCSVDVD 451
DB 413 NSATASSGDVLYIAGQPRYNTQVLIYRMEDGNKILQTLSEQIGSYFGSILTTTDD 472
QY 452 SNGSTDLVLIGAPHY-----YEQTR-GGQVSVCPPLPRQARWQCDVAVLY 495
DB 473 KDSNTDILLVGAAPVMYGTKEQKGVVYVALNQTREYQWLSLEPIKQTCSSRQNSCTT 532
QY 496 GEOGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHSO 554
DB 533 ENKNEPCGARFGTAIAVADNLDNDGNDIVIGAPLDDHGGVYIYHG-SGKTIRKEYAQ 591
QY 555 RIAGSKLSRLQVFGOSLGGQDLTMDGLVDLTGAGQHVLLRSOPVLRVKAIMEFNPR 614
DB 592 RIFSGGDKTLKFFGSGHGEMLNGDGLTDTVITGLGGALEFWSRDAVAVKVTWVFN 651
QY 615 EVARNVFECDQVVKQKAG--EVRVCLHVQ-KSTRDLRREGIOQSVVTDIALDSGRPH 671
DB 652 KVNIOKKNCH--MEGETVCINATVCFEVLKSKEDTIVEADLQ----YRVLDSLRQI 704
QY 672 SRAVENET-----KNSTRQVQLGTOTCETLKLQPNCLIEDPSPVILNLSVCT 725
DB 705 SRSFSGTQERKQVRNITVKSSC-----TKHSFYMLDKHDFQDSYR----ITLDFNLT-D 755
QY 726 PLSAFNLRPVLAEDAQRILFTALFPPEKNCNDNICDDLSITFSFMSLDCLVGVGPRE- 784
DB 756 PENG-----PVLDDSLPNSVHEIYPAKDCGNKEKICISLSLHVATEKDLIVRSQDK 810
QY 785 FNVTVTRNDGDSYRTQTFPPLDLSYRKVSTLQNRQSRQSWRLACASASTEVSGAL 844
DB 811 FNVSLTVKNTKDSAYNTRIVHYSPLNLFSGTEAIQKD-----SCESN----- 853
QY 845 KSTSCSINHPIPENSEVTFNITFDVDSKASLGN-KLLKANVTSENNMPTNKTEFQLE 903
DB 854 HNTIKGVGFPLRGRBMVTFKLFQNTSYLMBENVTIYLSATSDSEPPETLSDNVNITS 913
QY 904 LPVKYAVYVMTSHGVSYTKLNFPTASENTSRVMOHQVYSN-----LGORS-----L 950
DB 914 IPVKVEVGLQFYS-SASEHISIAANETVPEVINSTEDIGNEINIPYLIRKSGSPMPPL 972
QY 951 PISLVF-----LVPRNLQTVIMDRPQVTFSENLSSTCHTK-----RLPS 991
DB 973 KLSISFPNTSNGYPVLYPFLGLSS-----SENANCRPHIEDPFSINSKGKMTT 1021
QY 992 HSDFLAELRKAPVNCVSIACVORIQCDIPFFGIQE-----EFNATUK 1033
DB 1022 STD-----HLKRGITLDCNTCKFATITCNLTSSDISQVNVSLILWKPTFKYSFSSLNLTIR 1078
QY 1034 GNLSFDWYIKTSHNHLIIYSTAILPNDVSFTLLPGGAFVRSQETKVEPEVFNPLPL 1093
DB 1079 GEL-----RSENASLVLSNN-----QKRELAIQISKDGLPGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSSE 1128

DB 1115 WVLLSAPAGULLLLMLLILALWKIGFFKRPPLKKKMEK 1151
RESULT 10
ITAH HUMAN
ID ITAH_HUMAN STANDARD; PRT; 1189 AA.
AC Q9UKX5; Q9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITGAI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
RT human integrin alpha11 subunit (ITGAI1).";
RL Genomics 60:179-187(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kutsche-Gullberg M., Sejersen T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
RT integrin. A collagen-binding, i domain-containing, beta(1)-associated
RT integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
RP SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-CAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF109681; AAF01258.1; -;
DR EMBL; AF137378; AAD51919.2; -;
DR EMBL; AL359064; CAB94392.1; -;
DR HSSP; P17301; IAOX.
DR Genew; HGNC:6136; ITGAI1.
DR MIM; 604789; -;
DR GO; GO:0008305; C:integrin complex; TAS.


```
QY 998 ELRKAPVNCISIAVCRIQCIDIPFGIOEEFNATLKGNSLSPDWY-----IKTSHNHLIV 1052
Db 1052 LRRAPQLNHNSSVDVSNINCIRLVP-NQENFHLGLNL-----WLRSLKALAYKSKWIMVN 1107
QY 1053 STAEILFNDSTVFTLLPGOGAFVRSQTKVPEFVFN-----PLPLIVSGSSVGGLLLLA 1106
Db 1108 AALQROPH-SPF-----IFRBDPSRQIVFEISKQEDMQVPIIIVGSTLGGLLLLA 1158
QY 1107 LITAALYKLGPFK-RQYKDMMSGGPPGAEP 1136
Db 1159 LVLVALWKLGFRRSRRERE-----EGLDP 1183
```

RESULT 11

```
ITA2_BOVIN
ID ITA2_BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
GN ITGA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP MEDLINE=94193647; PubMed=7511592;
RA Kanata T., Puzon W., Takada Y.;
RT "Identification of putative ligand binding sites within I domain of
RL Integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
RL J. Biol. Chem. 269:9659-9663(1994).
CC !- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ.
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC !- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC !- SIMILARITY: Contains 1 VFMA domain.
CC !- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L25886; AAB59255.1;
CC PIR: I45914; I45914.
CC HSP: P17301; IAOX.
CC InterPro: IPR000413; Integrin_alpha.
CC Pfam: PF01839; FG-GAP; 3.
CC Pfam: PF00357; integrin_A; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS: PR01185; INTEGRINA.
CC PRINTS: PR00453; VWFADOMAIN.
CC SMART: SM00191; Int_alpha; 5.
CC SMART: SM00327; VWA_1.
CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE: PS0234; VWFMA; 1.
```

```
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER 1 18
FT SIGNAL <1 18
FT CHAIN 19 1170 INTEGRIN ALPHA-2.
FT DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1122 1143 POTENTIAL.
FT DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 92 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT ? ? VFMA.
FT DOMAIN 177 367 FG-GAP 3.
FT REPEAT ? ? FG-GAP 4.
FT REPEAT 423 475 FG-GAP 5.
FT REPEAT 477 538 FG-GAP 6.
FT REPEAT 540 599 FG-GAP 7.
FT REPEAT 604 656 FG-GAP 1.
FT CA_BIND 488 496 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 616 624 POTENTIAL.
FT SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1146 1150 GFPR MOTIF.
FT DISULFID 72 81 BY SIMILARITY.
FT DISULFID 669 726 BY SIMILARITY.
FT DISULFID 778 784 BY SIMILARITY.
FT DISULFID 854 865 BY SIMILARITY.
FT DISULFID 1008 1039 BY SIMILARITY.
FT DISULFID 1044 1049 BY SIMILARITY.
FT CARBOHYD 94 94 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 688 688 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1063 1063 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. ) (POTENTIAL).
FT VARIANT 580 580 G -> V.
FT VARIANT 588 588 R -> K.
FT VARIANT 725 725 R -> S.
SQ SEQUENCE 1170 AA; 128929 MW; EECEFC5F2448FB1 CRC64;

Query Match 18.5%; Score 1085; DB 1; Length 1170;
Best Local Similarity 27.6%; Pred. No. 2.2e-65;
Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFGQSVVOL---QGSRVVVGAPQEIIVANORGSLYQC--DYST 54
Db 19 YNVLGPKAKIPSGPSSQFGYAVQOFINPKGNLVLVSGSPWGGFPKRNMGDVYKCPVDUST 78
QY 55 GSCEPIRLQ-----VPVEAVNMNLSGLSLAATTSPPOLLACGPTVHQTCSNTVYVKGLC 107
Db 79 TTCEKLNQSTMSNSVNTKMTNMSLGLTRNVGTGGLTCGLPWAQCGSQYVTTGVC 138
QY 108 PLFGNSLRQPKQPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIETVMEQLK-- 165
Db 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVVCDESNSIYPWD--AVKNLEKFKVQGLDIG 194
QY 166 KSKTLFSLIMOYSEEFRIHFTKFPNNPNPSLIPKLTQLL---GRTHATGIRKVVRE 221
Db 195 PTKTOMGLIQYANNPRVFNLTFSKD---EMIKATSOQTFQYGGDLTNTFKAIQYARDT 251
QY 222 LFNITNGARKNAFKILITDGEKFGDPLGVEDVIPADREGVIYRIVGV-----GDAFR 276
Db 252 AYSTAAGRCPCATKVMVVVTDGESH-DGSKLKAVIDQCNDKNILRFGVILGYLNRNALD 310
QY 277 SEKSRQELNTVASKPRPDHVPQINNFALKTKTIONLREKIFAIETGTQSSSSFEHMSQ 336
Db 311 TKNLKEIKATIASIPTERHFFNVSDADLLEKAGTIGQIFSIETGVQG-GDNFQWMSQ 369
QY 337 EGFSAAIT--SNGPLSTVGSYDHWAGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
```

Db 370 VGFSAEYSPNNILMLGAVGAYDWSGTVVQKTPHGLIFSKQAFQIQLDRNHSSYLGY 429
Qy 392 AAIIILNRVQSLVGLAPRVQHIGLVAMFRONTGMESNANV-----KGTQIGAVFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNTVTVIOSQGDQIGSYFGSYL 484
Qy 446 CSVDVDSNGSTDLVLGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLGEQOPWG 503
Db 485 CAVDVNDKDTITDVLVGAPMYNDLKKEGRVYLTITKG--ILNWH--OFLSGPGLENA 541
Qy 504 RFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLTFHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFCSAIAALSDINMDCFNDFVIGVSPLENQNSGAVIYNGHEGM--IRLRYSQIKLSDRAF 600
Qy 562 SPLQVFGOSLSCGQDLTWGLVDLTGAGCHVILLRSQPVLRKAIMFNPREVARNVF 621
Db 601 SSHLOYFGRSLDGYGDLGSDITDVSAGFAGVQVWLWSQSIADVSVDASFTPKKI--TLL 658
Qy 622 ECNDQVVGKEAGEVAVLHVQKSTRDLREGQIQSVWVYDIALD----SGRPHSRAVFN 677
Db 659 NKVAEI-----KLKLCF-----SAKFRPTNQNNQVAIVYNTITIDEDQFSRVSIGLEPK 707
Qy 678 ETKNSTRROTQVGLGLTQTC--TLKLQLPNCIEDPVPVILRLNPSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMIVSQARCSEYIIHQEPS---DIISPLNLCMNISLENPGT-----756
Qy 734 RVLADDAORLFTALFPFKKNCNDNQCDDLSITP-----SPMSLCLVVGGRPRENVTV 789
Db 757 NPALAEAYSETKVFSIPFHKDCGDDGVCISDLVNLVQLPATQQQPFVSNQKRLTFSV 816
Qy 790 TVRNDGEDSVRTQVTFEFLDLSYRKVSTLQNRQSRQSWRLACESASST--EVSGLKSTS 848
Db 817 QLNKKESAYNTLIVDFSENLF-----ASMSMPVDGTEVTCQIASSQKSVT 864
Qy 849 CSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMMPTNKTEFQLELPVKY 908
Db 865 CNVGYPAKSKQOVTFITNFDFNLQ--NLQNAQASISPRALSQSEENMADNSVNLKSLLY 923
Qy 909 AVTMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQQR-----SLPISLVFLV 958
Db 924 DABIHIT-RSTNINFEVSLDGNVSSV--HSFE--DIGPKRIFISIKVTTGSPVMSA---976
Qy 959 PVRLNOTVWDRPQVTFSEN--LSSTCTHKE-----RLPSSHSDFLAE- 998
Db 977 -----SVIIHIIPOYTKDNPLMYLTCVHTDQAGDISCEABINPLKIGQTSVSSVFSKEN 1030
Qy 999 LRKAPVNCVSIACQICDIPFGIQEENFATLKGNSLFDWYIKTSHNHLIVSTAETI- 1057
Db 1031 FRHKLNCRTASCNSIMCWLRLDQVKGFLVNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNDVFTL-----LPGQAFVRSQTETKVBPFE--VPNPLPLIVGSSVCGLLLLALITA 1110
Db 1091 TYNPQIIVIBENTVTP-----LTIKPKHEKVEVPTGVIGSVIAGILLALLVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWLKLGFFKRYKDM 1155

RESULT 12
ITA2 MOUSE
ID ITA2 MOUSE STANDARD; PRT: 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M.; Chan B.M.; Uniyal S.; Onodera H.; Wang D.Z.;
RA Damjanovich L.; Latzer D.B.; Finberg R.W.; Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E.; Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha-
2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
C-PROPEPTIDES AND E-CADHERIN. MICE HOMOLOGOUS FOR A NULL MUTATION
IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29987; CAA82877.1; -;
DR EMBL; X75427; CAA53178.1; -;
DR PIR; S44142; S44142.
DR HSSP; P17301; LAOX.
DR MGD; MGI:96600; Icgaz.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178
FT DOMAIN 27 1129
FT TRANSMEM 1130 1151
FT DOMAIN 1152 1178
FT REPEAT 42 100
FT REPEAT ? 7
FT DOMAIN 185 375
FT REPEAT ? 2
FT REPEAT 431 483
FT REPEAT 485 546
FT REPEAT 548 607
FT REPEAT 612 664
FT CA_BIND 496 504
FT POTENTIAL.

Qy	680	KNSRTRQTVLGLTQTCET--LKQLPNCIEDVSPIVLRNLPSLVGCTPLPSAFGNLRPVL	737
Db	718	SRFLQKNVMVNEVQKSCSHHISIQKPS--DVMNPLDLKVDISLENPGTS-----PAL	768
Qy	738	AEDAQRLLFTALPFEKNCNGDNICQDDLST-----TFSPMSLDCLVVGPRENVT	788
Db	769	EAYSETVKVFSIPFYKECSGDGICISDLIDVOQLPAIQTSF-----IVSNQNKRLTFS	823
Qy	789	VTVNRDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRSMRLACESASST-EVSGALKST	847
Db	824	VILKNRGESAYNTIVLAEFSENLF-----ASFSPVDGTCEVSGSQSV	871
Qy	848	SCSINHPIIPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTE--FOLELP	905
Db	872	TCDVGYPALKSEQOQVTFINFDENLQ-NLQNOAAINQAFAPSEQ--ETNKADNSVSLTIP	928
Qy	906	VKYAVVMVYTSHGVS TKYLNLFASENTSRVMQHOYQVSNLQOR-----SLPISLV	955
Db	929	LLYDAELHLT-RSTNFINFEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSM	984
Qy	956	FLV-----PVRLNQTIVWDRPQVTF-SENLS	980
Db	985	LVTIHIPOYTKENKPNLLYLTGIOTDQAGDISCTAEINPLKPHTA---PSVSFKENFR	1040
Qy	981	STCHTKERLPSSHDFLAELRKAPVNCISIAVCQRIQCIPFGIQIEFNATLKGNLISFDW	1040
Db	1041	---HTKE-----LDCRTTSCSNITCWLKDLHMAEYFINVTVRVNRT	1080
Qy	1041	YIKTSHNHLIIVSTAELINDSVFTLLPGCAFVRSQTETKVPFEPVNPPLPIVGVSSVG	1100
Db	1081	FAASTFQTQVLTAAAEIDTHNPOLFVIEENAVTIPLIMKPKTEKAEVPT--GVIIGSII	1138
Qy	1101	GLLLALITAAALKYLGFFKQYKDM	1125
Db	1139	GILLLLAMTAGLWKLGFFKQYKDM	1163

RESULT 13

ITAZ_HUMAN

ID	ITAZ_HUMAN	STANDARD;	PRT;	1181	AA.
AC	P17301;				
DT	01-AUG-1990	(Rel. 15, Created)			
DT	01-AUG-1990	(Rel. 15, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)				
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).				
GN	ITGA2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NCBI	TaxID=9606;				
RN	[1]				
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.				
RC	TISSUE=Endothelial cells;				
RX	MEDLINE=89308879; PubMed=2545729;				
RA	Takada Y., Hemler M.E.;				
RT	"The primary structure of the VLA-2/collagen receptor alpha 2 subunit				
RT	(platelet GPIa): homology to other integrins and the presence of a				
RT	possible collagen-binding domain.";				
RL	J. Cell Biol. 109:397-407(1989).				
RN	[2]				
RN	SEQUENCE FROM N.A.				
RP	Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,				
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;				
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RN	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.				
RX	MEDLINE=98019223; PubMed=9353312;				
RA	Emeley J., King S.L., Bergelson J.M., Liddington R.C.;				
RT	"Crystal structure of the I domain from integrin alpha2betal.";				
RL	J. Biol. Chem. 272:28512-28517(1997).				
RN	[4]				
RP	VARIANT HPA-5 (BR)				

RX MEDLINE=94043762; PubMed=7901236;
RA Santoso S., Kalb R., Walke M., Kiefel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet allomorphs Br(a) and Brb are associated with a
RT single amino acid polymorphism in glycoprotein Ia (integrin subunit
RT alpha 2).";
RL J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=1074142;
RA Kroll H., Gardemann A., Fechter A., Haberbusch W., Santoso S.;
RA "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction.";
RL Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NATIT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1;
DR EMBL; AF512556; AAM34795.1;
DR PIR; A33998; A33998.
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MIM; 192974;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; P:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
KW SIGNAL 30 1181
FT CHAIN 30 1181
FT DOMAIN 1133 1132
FT TRANSMEM 1155 1154
FT DOMAIN 1155 1181
FT DOMAIN 1155 1161
FT REPEAT 45 103
FT REPEAT 2 7
FT REPEAT 188 378
FT DOMAIN 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT REPEAT 499 507
FT CA_BIND 563 571
FT CA_BIND 627 635
FT SITE 1157 1161
FT DISULFID 83 92
FT DISULFID 680 737
FT DISULFID 789 795
FT DISULFID 865 876
FT DISULFID 1019 1050
FT DISULFID 1055 1060
FT CARBOHYD 105 105
FT CARBOHYD 112 112
FT CARBOHYD 343 343
FT CARBOHYD 432 432
FT CARBOHYD 460 460
FT CARBOHYD 475 475
FT CARBOHYD 699 699
FT CARBOHYD 1057 1057
FT CARBOHYD 1074 1074
FT CARBOHYD 1081 1081
FT VARIANT 534 534
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT TURN 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT TURN 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
Query Match 18.2%; Score 1068; DB 1; Length 1181;
Best Local Similarity 26.9%; Pred. No. 3.1e-64;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Caps 43;

DR GO: GO:0007160; P:cell-matrix adhesion; TAS.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3-
DR Pfam: PF00092; VWA; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR04053; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 4.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; FALSE_NEG.
DR PROSITE: PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1167 INTEGRIN ALPHA-10.
FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1123 1145 POTENTIAL.
FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT 38 97 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT 167 350 VWFA.
FT REPEAT 365 427 FG-GAP 3.
FT REPEAT 428 482 FG-GAP 4.
FT REPEAT 483 545 FG-GAP 5.
FT REPEAT 546 605 FG-GAP 6.
FT REPEAT 608 660 FG-GAP 7.
FT DOMAIN 1134 1140 POLY-LEU.
FT CA_BIND 494 502 POTENTIAL.
FT CA_BIND 558 566 POTENTIAL.
FT CA_BIND 620 628 POTENTIAL.
FT DISULFID 76 86 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 681 736 BY SIMILARITY.
FT DISULFID 789 795 BY SIMILARITY.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT CONFLICT 844 844 I -> L (IN REF. 2).
FT FT CONFLICT 909 909 G -> V (IN REF. 2).
FT FT CONFLICT 926 926 E -> D (IN REF. 2).
SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AE0 CRC64;

Query Match 18.1%; Score 1064.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 5.3e-64;
Matches 353; Conservative 201; Mismatches 493; Indels 171; Gaps 44;

Qy 1 FNLDTENAMTFOENARG-FGOSVVOLOGSR---VVGCAPOEIVANORGSLYQC----- 50
Db 23 FNLDEHPRLFPFPPEAFEGYSVLQVHVGSGGQWMLVGPAMDGPGSDRGDVRCPVCGAH 82
Qy 51 -----DYSTG-SCPEIRLQVPEAVNMSGLSIAATTSPPQLLACGPTVHOTCSE 99
Db 83 NAPCAKHLGDYQIGNSHP-----AVNMHIGMSLLETDGGGFMACAPLWSRACGS 134
Qy 100 NTYVKGCLFLGSLNRQOPKQFPALRGCPQEDSDIAFLVDGSGSIIPHDPFRRAKEFIST 159
Db 135 SVFSSGICARVDASFQPGSLAPTAQR-CPTY-MDVVIVLDGSGNSIYP--WSEVQTFLR 190
Qy 160 VMEQL--KSKTFLSLMOYSEFRIHTFKFEFQNNPNRSLIKDITQLLGR-THTATGIR 216
Db 191 LVGKLFIDPEQIQVGLVQYGESPVHWSLGDGFRKKEEVVRAAKNLSRREGRETKTAQIM 250
Qy 217 KVVELEFNITGARKNAFKIILITDGEKF-GDPLGYEDVTPEADREGVIRYVIGV-GDA 274
Db 251 VACTEGFSQSHGGRPEARLLVWVVDGSHDGEELPAALKACEAGR--VTRYGIAVLGHY 308

Qy 275 FRSEKS-----RQELNTVASKPRDHVFOINNFEALKTIONOLREKI PAIEGTQTGSSSSSF 330
Db 309 LRRORDPSSFLREIRTIASDDERFFNVTDAAALTDIVDALGRIFGLESHANESSF 368
Qy 331 EHEMSQEGFSAAITSGNPLLSVTGSDYWAGGVFLYTSKEKSTFINMTVDS-----DMND 385
Db 369 GLEMSQIGFSTHRLKDGILFGWVGAYDWGGSVLMLEGGHRLFPFPRMALEDEFPPLQNH 428
Qy 386 AYLGYA-AAIILNRVQSLVLGAPRYQHIGLVAMFR-QNTGWESNANVKGTQIGAYGA 443
Db 429 AYLGSVSSMLLRGGRRLFLSGAPFRHRGKVIAPQLKKDGAVRVAQSQLOEQEIGSYFGS 488
Qy 444 SICSDVDVDSNGSTDVLVLIGAPHY--EOTRGQSVSCPLPRGORARWQCDVAVLYEQCQP 501
Db 489 ELCPDTRDGTDDVLLVAAPMFLGPQNKETGRVTVYLV--GQSSLLTLOGTQPEPPQD 546
Qy 502 MGRFGAALTVLGDVNGDKLTDVAIGAPGEEBDRGAYVLFHGTSGSGI SPHSQSRAGSKL 561
Db 547 -ARFGFAMGALPDLNQDGFADVAVGAPLEDGHOGALYLVHGTQ-SGVRRPHPAQRIAAASM 604
Qy 562 SPRLOVFGOSLSGGODLTMGDLVLTGCAQGHVLLLRSQPLRVKAIHFNPREVARNVF 621
Db 605 PHALSIFGRSVDRLDGLDGLDVAVGAQGAAILLSRPVHLTPSLEVTPOAISVWOR 664
Qy 622 ECNDQVVKGEAG--EVRVCLHVQKSTRDLREGQIQSVVYDLDLDSGRPHSRVAFNET 679
Db 665 DCCR---RQEAVALCLTAALCFQVTSRTQGRWDH---QFYMRFTASLDEWTAGARAFDGS 718
Qy 680 --KNSTRQTVLGLTQTCETLKLQPNICIEDVSPVILRLNFSLVGTPLSAGNLRPVL 737
Db 719 GQRLSPRLRLSVG-NVTCQLHFHVL-DSDYLRPVALTVTFALDNTTKPG-----PVL 771
Qy 738 ABDAQLFTALPFFKNCNNDNI CQDLSITFSFMSLDC-----LVVGGPRFNVTV 789
Db 772 NEGSPTSIQKLVFPSKDCPDNECVTDLVQ---VNMDIRGSRKAPFVVRGGRKRVLVST 828
Qy 790 TVRNDGEDSVRTQVTFPFLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKSTSC 849
Db 829 TLNKRKENAYNTSLSIIFSRNL---HLASLTQOR-ESPIKVECAAPSA-----HARLC 877
Qy 850 SINHPFPENSEVTNITFDVDSKASLG---NKLL-----LKANVTSEKNMPTNKTEFO 901
Db 878 SVGHVPVQTAQVTFLEFEFSCSSLLSQVFGKLTASSDSLSRNGTLQENTAQT----- 931
Qy 902 LELPVKAYVMVTSHGVTSKYLNFTASENTSRVMOHQY-----QVSNLG 946
Db 932 -----SAYIQVEPH-----LLFSESTLHRYEVHYGTLPVGPGEFKTLRLVQNLG 978
Qy 947 ---QRSPLPSLVLPV-----VRLAQTVIWDPRQVTFSENLSSSTCHTKERLPSSH 994
Db 979 CVYVSGLIISA--LLPAVAHGNYFLSLSOVI-----TNNASCIQNLTEPPGPPV 1027
Qy 995 FLAELRKAPVWNCIAVCQRIQCDIPFGFIQEBFNATLKGNSLFDWYIKTSHNHLIVST 1054
Db 1028 HPELOHTVRLNGSNQCQVVRCHLQOLAKGTEVSVGLLRVLVHNEFFRRAKFSLTWVST 1087
Qy 1055 AEILFVNDSVFTLLPGOGAFVRSQTEYKVPFVFNPLPLIVGSSVGLLLALLITAAIYK 1114
Db 1088 FELGTBEGSVLQLTEASRWSESLLEV-VQTRPILISLWILIGSVLGLLLALLVFLCLWK 1146
Qy 1115 LGFF-----KROYK 1123
Db 1147 LGFFAHKKIPEEKREEK 1164

RESULT 15
ITAL_RAT
ID ITAL_RAT STANDARD; PRT; 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1) (CD49a).

GN ITGA1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90338125; PubMed=2380249;

RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A., Esch F., Carbonetto S., Reichardt L.F.;

RA "Molecular cloning of the rat integrin alpha 1-subunit: a receptor for laminin and collagen."

RT J. Cell Biol. 111:709-720(1990).

RL [2]

RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.

RX MEDLINE=99313197; PubMed=1038626;

RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliansky V., Gotwale P.J., Karpusas M.;

RA "Crystal structure of the alpha1beta1 integrin I-domain: insights into integrin I-domain function."

RL FEBS Lett. 452:379-385(1999).

CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN COLLAGEN.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1 ASSOCIATES WITH BETA-1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X52140; CRA36384.1; .

DR PIR; A35854; A35854.

DR PDB; 1CK4; 03-MAY-00.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWFA.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 5.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS02034; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; Repeat; Calcium; Magnesium; 3D-structure.

FT SIGNAL 1 28

FT CHAIN 29 1180

FT DOMAIN 29 1142

FT TRANSMEM 1143 1165

FT DOMAIN 1166 1180

FT REPEAT 44 103

FT REPEAT ? ?

FT DOMAIN 175 388

FT REPEAT 377 432

FT REPEAT 433 484

FT REPEAT 485 565

FT REPEAT 567 626

FT REPEAT 629 681

FT CA_BIND 497 505

POTENTIAL.

INTEGRIN ALPHA-1.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

FG-GAP 1.

FG-GAP 2.

VWFA.

FG-GAP 3.

FG-GAP 4.

FG-GAP 5.

FG-GAP 6.

FG-GAP 7.

POTENTIAL.

FT CA_BIND 579 587 POTENTIAL.

FT CA_BIND 641 649 POTENTIAL.

FT SITE 1168 1172 GFFKR MOTIF.

FT DISULFID 82 92 BY SIMILARITY.

FT DISULFID 687 696 BY SIMILARITY.

FT DISULFID 702 755 BY SIMILARITY.

FT DISULFID 807 813 BY SIMILARITY.

FT DISULFID 877 885 BY SIMILARITY.

FT DISULFID 1029 1062 BY SIMILARITY.

FT DISULFID 1066 1073 BY SIMILARITY.

FT CARBOHYD 100 100 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 105 105 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 112 112 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 217 217 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 317 317 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 341 341 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 402 402 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 418 418 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 531 531 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 698 698 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 747 747 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 779 779 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 820 820 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 839 839 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 882 882 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 907 907 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 938 938 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 965 965 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 973 973 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 1007 1007 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 1084 1084 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 1103 1103 N-LINKED (GLCNAC.) (POTENTIAL).

FT CARBOHYD 1114 1114 N-LINKED (GLCNAC.) (POTENTIAL).

SQ SEQUENCE 1180 AA; 130808 MW; 8E5DA2B2E02362EE4 CRC64;

Query Match

18.1%; Score 1062; DB 1; Length 1180;

Best Local Similarity 27.4%; Pred. No. 8e-64;

Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

Qy 1 FNLDTENAMTFQENARG-FQGSVVQL---QGSRVVVCAPQEIIVANORGSLYQCDYGTGS 56

Db 29 FNVDVKNMSFSGFVEDMFQYVQYENEEGKWLIGSLVGPQKARTGDVYKCPVGRER 88

Qy 57 CEP-IRLOVPVEA-----VNMSLGLSLAATTSPPQLLACQPTVHQTCSEYVYKGL 106

Db 89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGGFLACGLVYRCHLHYTTGI 147

Qy 107 CFLFGSNLRQOPKFPALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK- 165

Db 148 CSDVSPTFQVNSFAP--VQECSTQ-LDIVLDGSGNSIYP--WESVIAFLNDLLKRMDI 202

Qy 166 -KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLLG-RTHATGIRKVKVREL 223

Db 203 GPKQTQVGIQVYGENVTHEFNLNKYSSTEEVLVAANKIGRGGGLQTMALGIDTARKEAF 262

Qy 224 NITNGARKNAFKILITDCEKFGDPLGYEDVTPADREGVIRVVGVDADR-----SE 278

Db 263 TEARGARRGVKKVMVIVTDGESH-DNYRLKQVTDCEDENIQRFSAIILGHYNGNISTE 321

Qy 279 KSRQELNVASKPPRPHVFOINNFEALKTQNLQRKIFAIEGTQTQSSSSFFHEMSQEG 338

Db 322 KFVEEIKSIASEPTEKHFFNVSDDELALVTIVKALGERIFALEATADOSAASFEMEQTG 381

Qy 339 FSAAITSNGLPLSTVGSYDWAGGVFLYTSKEKSTFFINMT--RVDSDND---AYLGAAAA 393

Db 382 FSAHYSQDWMLGAVGAYDNNGTVVMQANQMVIPIHNTTFTQTEPAKNEPLASLYGTVN 441

Qy 394 IILNRVQSLVLGAPRYQHIGLVAMFRONTGMESNANVKGTQIGAVFGSLCSVDVDSN 453

Db 442 SATIPGDVLYIAGOPRYNHTGVVIYKMEGDNINILQTLGGEQIGSVFGLTTIDDKD 501

Qy 454 GSTDLVLIGAPHY-----YEQTR--GGQVSVCPPLRQGRARWQCDAVLGE 497

```
Db 502 SYTDLLLVGAPWMTGTEKEEQKVYVYAVNQTFFEQMSLEPIRQTCCSSSLKDNSTKEN 561
Qy 498 QGQPMG-RFGAALTUVGDNVNGKLTVAIGAGEEDNRGAVYLFHGTSGSGISPSHSQRI 556
Db 562 KNEPCGARFCTAAAVKDLNVDGFDNDWI GAPLEDDHAGAVIYHG-SGKTIREAYAORI 620
Qy 557 AGSKLSPRLQYFGOSLGGODLTMDGLVDTLVGAQGHVLLLRSPVLRVKAIMEFNPREV 616
Db 621 PSGDGKTLKFFQOSIHGEMDNGDGLTDTTIGLGAALFWARDVAVVKTWTFNPFNKV 680
Qy 617 ARNVFECNDQVVKGEAG--EVRAVCLHVQ-KSTRDLREGIQSVVTYDLALDSGRPHSR 673
Db 681 NIQKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ----YRVTLDLSLRQISR 733
Qy 674 AVFNET-----KNSTRQTVLGLTQTCETLKLQLPNCI-----EDVSPVIVLRL 718
Db 734 SFESGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVRVTL 776
Qy 719 NFSLVGTPLSAFCNLRVLAEDAQRFLTALPFEKNCNDNICQDDLSITFFSMJLDCIV 778
Db 777 DFNLT-DPENG-----PVLDDALPNSVHEHIPPAKOCGNKERCISDLTLNVSTTEKSLI 930
Qy 779 VGGPRE-FNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRSWLACESASS 837
Db 831 VKSQHDKFNVLTVKNKGSAYNTRTVQHSNPLIFSGIEEIQKD-----SCSN-- 880
Qy 838 TEVSGALKSTCSINPIPPENSEVTNITFDVDSKASLGKLL-LKANVTSENNMPTN 896
Db 881 -----QNTICRVGYPLFRAGETVTKIIFQNTSHLSENAIHLATSDSSEPLESLN 933
Qy 897 KTEFQLELPVKYAV---YMWVTSHGVS-----KYNFTASENTRVMQHQYQVSNL 945
Db 934 DNEVNISIPVKEVGLFYSSASEHHISVAANETIPEFINST--EDIGNEINVFTYIRK 991
Qy 946 GORSPL---ISLVF-----LVPVRLNQTVID-----RP-----Q 972
Db 992 GHFPMPELQISFPNLTADGYPVLYPIG-----WSSSDNVNCRPSRLEDPFGINSKK 1045
Qy 973 VTFS-----ENUSSTCHTKERIPSHSDFLAELURKAPVNVCSIAVCQRIQCDI--- 1019
Db 1046 MTISKSEVLKRGTIQCSSTC-----GVATITCSLLPSDLSQVNVSL 1088
Qy 1020 ---PFGIQEEF---NATLGNLSFDWYIKTSHNHLIIVSTAILFNDSVFTLLPGOGAF 1073
Db 1089 LWKPTF-IRAHFSSNLNITLRELK-----SENSLTSSN----- 1123
Qy 1074 VRSQTEKVPFEPNPLPL--IVGSSVGGLLLALITAAALYKLGFFKQYKDMMSSE 1128
Db 1124 RKRELAIQISKDGLPGVPLWVILLSAFAGULLLMLLALWKIGFFKRLPKKKMEK 1180
```

Search completed: November 25, 2003, 14:17:34
Job time : 13.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirts:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4381	74.6	1151	11	Q9J130
2	3901.5	66.4	1036	11	Q8CA73
3	3809.5	64.8	920	6	Q28984
4	3480	59.2	1169	4	Q81VA6
5	3320.5	56.5	1169	11	Q9QXH4
6	3244.5	55.2	1161	11	Q9QYE7
7	1534.5	26.1	1161	11	Q9RTV4
8	1524	25.9	1160	11	Q9R200
9	1409	24.0	1196	13	Q9RTF1
10	1359.5	23.1	1086	4	Q96HB1
11	1358.5	23.1	1187	13	Q98TF0
12	1278	21.7	927	6	Q8HZV0
13	1167.5	19.9	1167	11	Q88340
14	1123	19.1	1167	11	Q88341
15	1053	17.9	1171	13	Q42094
16	1045	17.8	1038	11	Q8BS01

17	1018.5	17.3	1160	6	Q8MKF4
18	1018	17.3	895	11	Q9WUF8
19	865	14.7	348	4	Q8TES5
20	850	14.5	1332	5	Q8BPQ8
21	808	13.8	205	11	Q63001
22	753.5	12.8	780	13	Q06271
23	738	12.6	823	4	Q8WY18
24	686.5	11.7	823	11	Q8CB84
25	669	11.4	1032	11	Q61989
26	643	10.9	1036	11	Q91YD5
27	642.5	10.9	1033	6	Q9BGU3
28	623.5	10.6	257	11	Q8C270
29	619.5	10.5	1041	5	Q9UB90
30	614.5	10.5	1041	5	Q76378
31	578.5	9.9	1054	5	Q9U6S1
32	578.5	9.8	1034	13	Q98TT7
33	550	9.4	1033	13	Q42598
34	534	9.1	1016	13	Q91779
35	530	9.0	974	11	Q924W2
36	529	9.0	1073	11	Q8CC06
37	526	9.0	1047	6	Q9MZD6
38	525.5	8.9	1007	6	Q9GK48
39	512.5	8.7	1034	6	Q9TUN4
40	509.5	8.7	1036	6	Q9TUN6
41	506.5	8.6	1012	11	Q70304
42	491.5	8.4	1049	5	Q8SY51
43	475	8.1	833	5	Q9BPQ7
44	469.5	8.0	1036	6	Q9TU44
45	468.5	8.0	604	11	Q8BQ25

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
 ID Q9J130
 AC Q9J130;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr., Zeria K. Jr.;
 RT "Cloning of the rat CD11b cDNA sequence."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268593; AAF81280.1; -;
 DR HSSP; P11215; 1BHQ.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5_
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWF_A; 1.
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 11; Length 1151;
 Best Local Similarity 72.7%; Pred. No. 1.6e-316;
 Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;
 Oy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAANQRGSLVQCQDSTGSCPEI 60
 |||||
 Db 17 FNLDTENPMTFQENASFGQSVIQLGETRVVVVAAQEVKAVNQIYQCDYSTNRCDPI 76
 |||||

```
QY 61 RLOVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 120
DB 61 RLOVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 120
QY 77 PLQVPPEAVNMSLGLSLAATTSPPQLACGPTVHQNCKENTYVYKGLCYLFGSNLRKPKQ 136
DB 77 PLQVPPEAVNMSLGLSLAATTSPPQLACGPTVHQNCKENTYVYKGLCYLFGSNLRKPKQ 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEOLKSKTLFSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEOLKSKTLFSLMOYSEEF 180
QY 137 FPEALRGCPQESNIAFLIDGSGSINTIDFQKMKFEVSTVMDQFQSKTLFSLMOYSDFF 196
DB 137 FPEALRGCPQESNIAFLIDGSGSINTIDFQKMKFEVSTVMDQFQSKTLFSLMOYSDFF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240
DB 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240
QY 197 RTHFTFDFKRNPKSHVRPIRLNGRTKTASGIRKVVRELFNITNGARKNAFKLIL 256
DB 197 RTHFTFDFKRNPKSHVRPIRLNGRTKTASGIRKVVRELFNITNGARKNAFKLIL 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 300
DB 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 300
QY 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 316
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 316
QY 301 NFPAALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 360
DB 301 NFPAALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 360
QY 317 NFPAALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 376
DB 317 NFPAALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
QY 377 GAFLYPSKDKASPIINTRIDSDMNDAYLGYASAVISNRVQSLVGLGAPRYQHIGLVAMFR 436
DB 377 GAFLYPSKDKASPIINTRIDSDMNDAYLGYASAVISNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 480
DB 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 480
QY 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 496
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 496
QY 481 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
DB 481 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
QY 497 PRG-RARWQCEALHGDQHPWRBFGNALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 555
DB 497 PRG-RARWQCEALHGDQHPWRBFGNALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 600
QY 556 HGASVASISTPHSQRITAGARFSGQLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 615
DB 556 HGASVASISTPHSQRITAGARFSGQLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 615
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 660
DB 601 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 660
QY 616 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 675
DB 616 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 675
QY 661 YDLALDSGRPHSRAVFNETKNSRTRQTLGLTQTCETLKLQPCIEDPVSIVLRNLF 720
DB 661 YDLALDSGRPHSRAVFNETKNSRTRQTLGLTQTCETLKLQPCIEDPVSIVLRNLF 720
QY 676 YDLALDGRSVTRAFFDETNGILRRIRVFLGTQKCTELKLIIPDCVDNSVSIILRLNY 735
DB 676 YDLALDGRSVTRAFFDETNGILRRIRVFLGTQKCTELKLIIPDCVDNSVSIILRLNY 735
QY 721 SLVGTPLSAGNLRLPVLAEADAQRLFTALFPFKKNCNDNTICQDLSITFSMLDCLVWG 780
DB 721 SLVGTPLSAGNLRLPVLAEADAQRLFTALFPFKKNCNDNTICQDLSITFSMLDCLVWG 780
QY 736 TLVGEPLRSSRDLRLPVLAEADAQRLFTALFPFKKNCNDNTICQDLSITFSMLDCLVWG 795
DB 736 TLVGEPLRSSRDLRLPVLAEADAQRLFTALFPFKKNCNDNTICQDLSITFSMLDCLVWG 795
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVYKVVSTLONORSQSRWLACASSTEV 840
DB 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVYKVVSTLONORSQSRWLACASSTEV 840
QY 796 DSRDFDVSVTLRNDGDSYRTQVTFPPFLDLVYKVVSTLONORSQSRWLACASSTEV 854
DB 796 DSRDFDVSVTLRNDGDSYRTQVTFPPFLDLVYKVVSTLONORSQSRWLACASSTEV 854
QY 841 SGALKSTCSGINHPIPFENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
DB 841 SGALKSTCSGINHPIPFENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
QY 855 QGVKSTIWDINHPIPFANSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 914
DB 855 QGVKSTIWDINHPIPFANSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 914
QY 901 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSLPIISLVFLVVP 960
DB 901 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSLPIISLVFLVVP 960
QY 915 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSLPIISLVFLVVP 974
DB 915 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSLPIISLVFLVVP 974
QY 961 RLNQTIVDRPOVTFSENLSSTCHTKERLPSHDSFLAELKAPVNVCSIAVCQRIQCDIP 1020
DB 961 RLNQTIVDRPOVTFSENLSSTCHTKERLPSHDSFLAELKAPVNVCSIAVCQRIQCDIP 1020
QY 975 QINKVTIWDIPQVTFSENLSSTCHTKERLPSHDSFLAELKAPVNVCSIAVCQRIQCDIP 1034
DB 975 QINKVTIWDIPQVTFSENLSSTCHTKERLPSHDSFLAELKAPVNVCSIAVCQRIQCDIP 1034
QY 1021 FFGQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSTFLLPGOGAFVRSOTET 1080
DB 1021 FFGQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSTFLLPGOGAFVRSOTET 1080
QY 1035 SFNSKEIFNVTLOGLNLLFDWYIETSHDHLVSTAEILFNDVSTFLLPGOGAFVRSOTET 1094
DB 1035 SFNSKEIFNVTLOGLNLLFDWYIETSHDHLVSTAEILFNDVSTFLLPGOGAFVRSOTET 1094
QY 1081 KVEPFEPNPLITVSGSVGLLALLITALYKLGFFKRYQKMDMSEGGPPGAP 1137
DB 1081 KVEPFEPNPLITVSGSVGLLALLITALYKLGFFKRYQKMDMSEGGPPGAP 1137
QY 1095 KVEPFEPNPLITVSGSVGLLALLITALYKLGFFKRYQKMDMSEGGPPGAP 1151
DB 1095 KVEPFEPNPLITVSGSVGLLALLITALYKLGFFKRYQKMDMSEGGPPGAP 1151
```

```
RESULT 2
QBCA73 ID Q8CA73 PRELIMINARY; PRT; 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; -.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;

Query Match 66.4%; Score 3901.5; DB 11; Length 1036;
Best Local Similarity 66.1%; Pred. No. 6.1e-281;
Matches 752; Conservative 125; Mismatches 142; Indels 119; Gaps 2;

QY 1 FNLDTENAMTTFQENARGFQGSVVQLQGSVVVGAPOEIVAAHQRSGLYQCDYSTGSCPEI 60
DB 17 FNLDTEHPMTFOENAKGFQGSVVQLGTSVVVAAPOEAKAVQGTALYQCDYSTSRCHPI 76
QY 61 RLOVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVYKGLCYLFGSNLRPPOQ 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEOLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQESNIAFLIDGSGSINTIDFQKMKFEVSTVMDQFQSKTLFSLMOYSDFF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240
DB 197 RHFTFDFKRNPKSHVRPIRLNGRTKTASGIRKVVRELFNITNGARKNAFKLIL 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 300
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVFOIN 316
QY 301 NFEALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 360
DB 317 NFEALNTIONQREKIFALEGTQTSSEFHEMSOEGFSAATITSGPLLSVGVSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGYASAVILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 480
DB 437 ENFGTWEPTSING----- 450
QY 481 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
DB 451 ----- 450
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 600
DB 451 -----SQRIGAHFSPGLQYFGQSLGGQDLTMDGLVDLTVAIGAPCEDNRGAVYLF 498
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 660
DB 499 PVLVRVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLRREGIQSVVT 558
```



```
QY 661 YDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILRLNF 720
D 661 YDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILRLNF 720
D 661 YDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILRLNF 720
QY 721 SLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCVLVW 780
D 721 SLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCVLVW 780
D 721 SLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCVLVW 780
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRL-ACSSASSTE 839
D 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRL-ACSSASSTE 839
D 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRL-ACSSASSTE 839
QY 840 VSGALKSTCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNKTE 899
D 840 VSGALKSTCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNKTE 899
D 840 VSGALKSTCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNKTE 899
QY 900 FOELPLPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVLVLP 959
D 900 FOELPLPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVLVLP 959
D 900 FOELPLPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVLVLP 959
QY 960 VRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIOCDI 1019
D 960 VRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIOCDI 1019
D 960 VRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIOCDI 1019
QY 1020 PFFGIQEBFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDSTFTLLPGQGFVRSQTE 1079
D 1020 PFFGIQEBFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDSTFTLLPGQGFVRSQTE 1079
D 1020 PFFGIQEBFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDSTFTLLPGQGFVRSQTE 1079
QY 1080 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1137
D 1080 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1137
D 1080 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1137
QY 979 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1036
D 979 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1036
D 979 TKVEPFEVNPPLIVGSSVGLLLALITIAALYKLGFFKQYQKDMMEGPPGAPQ 1036

RESULT 3
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSP; F11215; I8HQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS0234; VWF A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.8%; Score 3809.5; DB 6; Length 920;
Best Local Similarity 78.9%; Pred. No. 3.5e-274;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;

QY 118 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 177
D 118 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 177
D 118 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 177
D 1 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 60
D 1 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 60
D 1 POKPEALRGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMYS 60
```

```
QY 178 EEFRIHFTFKFQNNPNRSLIKPITQLLQRTHTATGIRKVVRELFNITNGARKNAFKIL 237
D 178 EEFRIHFTFKFQNNPNRSLIKPITQLLQRTHTATGIRKVVRELFNITNGARKNAFKIL 237
D 178 EEFRIHFTFKFQNNPNRSLIKPITQLLQRTHTATGIRKVVRELFNITNGARKNAFKIL 237
QY 238 ILITDGSKFGDPLYEDVIVPEADREGVIRVIVGVDAFRSEKSRQELNITVASKPRDHVF 297
D 238 ILITDGSKFGDPLYEDVIVPEADREGVIRVIVGVDAFRSEKSRQELNITVASKPRDHVF 297
D 238 ILITDGSKFGDPLYEDVIVPEADREGVIRVIVGVDAFRSEKSRQELNITVASKPRDHVF 297
QY 298 QINNPEALKTIQNLREKIPAEIQTOTQSSSSPHEMSQEGFSAAITNSGPLLSTVGSYD 357
D 298 QINNPEALKTIQNLREKIPAEIQTOTQSSSSPHEMSQEGFSAAITNSGPLLSTVGSYD 357
D 298 QINNPEALKTIQNLREKIPAEIQTOTQSSSSPHEMSQEGFSAAITNSGPLLSTVGSYD 357
QY 358 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
D 358 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
D 358 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
QY 417 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
D 417 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
D 417 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRNVQSLVGLGAPRYOHIGLVA 417
QY 478 CPLPRGORARWQCDALVLYGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
D 478 CPLPRGORARWQCDALVLYGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
D 478 CPLPRGORARWQCDALVLYGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
QY 538 YLPHGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLL 597
D 538 YLPHGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLL 597
D 538 YLPHGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLL 597
QY 598 RSQVPLRVKAIMENPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
D 598 RSQVPLRVKAIMENPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
D 598 RSQVPLRVKAIMENPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
QY 658 VVTVDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILR 717
D 658 VVTVDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILR 717
D 658 VVTVDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQLCNIBDPSPVILR 717
QY 718 INFSLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDC 777
D 718 INFSLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDC 777
D 718 INFSLVGTPLSARGLRPLVLAEDQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDC 777
QY 778 VVGPREPNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRLACESASS 837
D 778 VVGPREPNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRLACESASS 837
D 778 VVGPREPNVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRORSWRLACESASS 837
QY 838 TEVSGLKSTSCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNK 897
D 838 TEVSGLKSTSCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNK 897
D 838 TEVSGLKSTSCSINHPFIPENSEVTNITFDVDSKASGLNKLKLLKANVTSENMPRTNK 897
QY 898 TEFQLELPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVL 957
D 898 TEFQLELPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVL 957
D 898 TEFQLELPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVL 957
QY 958 VPVRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIO 1017
D 958 VPVRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIO 1017
D 958 VPVRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIO 1017
QY 1018 DIPFGIOEENATLKGNSL 1038
D 1018 DIPFGIOEENATLKGNSL 1038
D 1018 DIPFGIOEENATLKGNSL 1038
QY 900 DIPFGIOEELKVLKGNLSF 920
D 900 DIPFGIOEELKVLKGNLSF 920
D 900 DIPFGIOEELKVLKGNLSF 920

RESULT 4
Q81V46 PRELIMINARY; PRT; 1169 AA.
ID Q81V46
AC Q81V46;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237.1;
SQ SEQUENCE 1169 AA; 128521 MW; 17B484FEFC79EB6 CRC64;

Query Match 59.2%; Score 3480; DB 4; Length 1169;
Best Local Similarity 61.2%; Pred. No. 1.7e-249;
Matches 691; Conservative 141; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGCQSVVQLQSGSRVVVGAPOEIVAAQNRGSLYOCDSYSGCEPI 60
DB 20 FNLDTBELTAFAFRVDSAGFSDGVVQYANVSWVGAPOKITAANQOTGLYQCGYSTGACEPI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK 120
DB 80 GLQVPEAVNMSLGLSLAATTSPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIPTVMOLKSKTFLSLMOYSEEF 180
DB 138 LPVSRQCEPQEQDIFVLIDGSGSISSRNFAFMNEFRAVISQFORPSTQFSLMOPSNKF 197

QY 181 RIHFTKPEQNPNPSLIKPIITOLGRTHATGIRKVVRELPNITNGARKNAFKLILI 240
DB 198 QTHFTFEFRSRNPSLLASVHQLOGFTYTATAIQNVVHRUFHASYGARRDAAKTLIVI 257

QY 241 TDCEKFGDPLGYEDVTPADREGVIRYVGVGDAFRSEKSRQELNVTASKPRDRHVFQIN 300
DB 258 TDCKKSGDSLVDKVDVPMADAAGIIRYAGVGLAFQNRNSKELNDIASKPSQEHIFKVE 317

QY 301 NFPAKTIQNQLREKIFAIEGTQTSSSSFEHMSQEGFSAATITSGPLLLSTVGSYDWA 360
DB 318 DFDALQIDQNLKEKIFAIEGTETSSSFELEMAQEGFSAVTPDGPVLGAVGSGTWSG 377

QY 361 GVFLYTSKESKSTINTRVDSNMNDAYLGAAAILLRNVSQSLVLCAPRYOHLGLVAMER 420
DB 378 GAFLYPPNMSPTFINNSQENMDRDSYLGYSTELALWKGVQSLVLCAPRYOHTGKAVIF 437

QY 421 QNTGMNESANVAGTQIGAFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQGVSCPL 480
DB 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQGVSCPL 497

QY 481 PRGORARWCDVLYGEGQGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
DB 498 PRGWR-RWVCDVLYGEGQGWGRFGAALTVLGDVNGDKLTDVVI GAPGEKENRGAVLYF 556

QY 541 HGTSGSGISPSHSORAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLILRSQ 600
DB 557 HGVLPSPISPSHSQRIAGSQLSRSLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLR 616

QY 601 PVLVRKAIMEFNPVARNVFNCDQVVKGEAGEVRVCLHVOKSTRDLRLEGQIQSVVT 660
DB 617 PVLWVGSMQFIIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLLGSDLOSV 676

QY 661 YDIALDSGRPHSRVAFNETKSNTRQTVLGLTQTCETLLQOLPNCIEDPVSPVLRNLN 720
DB 677 LDIALDPGRLSPRATFQETKNRSLRVRVLGLKAHCENFNLLPSCVEDSVTPIITLRNL 736

QY 721 SLVGTPLSAGNLRPLVLAEDAORLFTALPFPEKNCNDNTCODDLSITTSFMSLDCLVG 780
DB 737 TLVGKPLAFARNRPLMLAADAQRYFTASLPFEKNCAGADHICQNLGISFSFPGKLLLV 796

QY 781 GPREFNVTVVRNDGDSYQTVTFPPPLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 840
DB 797 SNLELNAEVMWVNDGDSYGTVTTFSHPAGLSYRYAEGQKQQLSLHLTCDSPAVG-- 854

QY 841 SGALKSTSCINHIPFPENSEVTFNITFDVDSKASLGNKULLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCINHLIFGGAQIITFLATFDVSPRAVLGDRLLLTANVSENNTPTSTKTF 914

QY 901 QLELPVKYAVYVTVVSSHEQFTKYLNFSESEKESHVAMHRQVNNLQORDLPVSNFNP 974
DB 960 VRLNQTVIWRDPQVTFSENLSSTCHTRELPSHSDFLAELKAPVWNCISIAVCORIQCDI 1019
DB 975 VELNQAEMVMDVEVSHFQNPSPSRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034
QY 1020 PFFGIQBEFNATLKNLSFDWYIKTSHNHLIIVTAETILFNDVSFTLLPGGAPVRSQTE 1079
DB 1035 PSFSVQBELDTLTKNLSFGWVRQLQKXSVSWVAELFTDTSVYSQLPGQEAEMRAQTT 1094
QY 1080 TKVPEFPEVNPPLIVGVSSVGLLLALITAAALYKLGFFKRYQKDMSE 1128
DB 1095 TVLEKYKVNPTPLIVGSSIGLLALLITAVLYKVGFFKRYKEMEE 1143

RESULT 5
QYQH4
ID QYQH4 PRELIMINARY; PRT; 1169 AA.
AC QYQH4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAF23492.1;
DR HSSP; P11215; 1BHQ.
DR MGD; MGI:96609; Itgax.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.5%; Score 3320.5; DB 11; Length 1169;
Best Local Similarity 56.8%; Pred. No. 1.3e-237;
Matches 647; Conservative 172; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFOENARGCQSVVQLQSGSRVVVGAPOEIVAAQNRGSLYOCDSYSGCEPI 60
DB 20 FNLDAEKLTHFMDGAEGFHSVLQYDSSVWVVGAPKEIKATNIQIGLYKCYHTGNCPEI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK 120
DB 80 SLQVPEAVNMSLGLSLAATTSPQLLACGPTVHHTCTRENIVLTGLCFLLSLSPKQS-QN 138

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIPTVMOLKSKTFLSLMOYSEEF 180
DB 139 FTAQOCECPQDQDIFVLIDGSGSISSDTDFERKMLDFVKAVMSQLQRPSTFSLMQPSDYF 198

QY 181 RIHFTKPEQNPNPSLIKPIITOLGRTHATGIRKVVRELPNITNGARKNAFKLILI 240
DB 199 RVHFTFNFNISTSSPLSLGSRQLRGYTYTASAKHVITELFTTQSGARODATKVLIVI 258

QY 241 TDCEKFGDPLGYEDVTPADREGVIRYVGVGDAFRSEKSRQELNVTASKPRDRHVFQIN 300
```

```
Db 259 TDGRKQGNLSYSDVPIWAEASIRYAIGVKAFYNEHSHSQELKATASMPSEHYVSFE 318
Qy 301 NFEALKTIONOLREKI FAIEGTOTGSSSSFEHMSOEGFSAITNSGFLLTGVSQYDAG 360
Db 319 NFDALKDLENQLEKI FAIEGTETPSSSTFEHMSOEGFSAVFPDGPVLGAVGSFWSG 378
Qy 361 GVFLYSKEKSTFIMNTVDSMDNDAYLGAAAAIILNRVOSLVGAPRYQHIGLVAMFR 420
Db 379 GAFLYPSNMRTPTFINMSQENEDMRDYLGYSTALAFWKGVHSLILGAPRHQHTGKVIPT 438
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQVSCPL 480
Db 439 QESRHWPKSEVRGTQIGSYFGASLCSVDMDRDSGLVLIGVPHYYEHTGGQVSCPM 498
Qy 481 PRGQARWOCDAVLGEQGPWGRFGAALTVLGVNGDKLTVAIGAPGEDNKGAVYLF 540
Db 499 P-GVGSRWHCGTTLHGEQHPWGRFGAALTVLGVNGDSLADVAIGAPGEENKGAVYIF 557
Qy 541 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPQRISASQIPSRIQYFGQSLSGGQDLTRDGLVDLAVGSKRVLRLTR 617
Qy 601 PVLRYKAIWENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVT 660
Db 618 PILRVSPVTHETPABISRSVFECQVAPQETLSDATVCLHIHESPKTQL--GDLRSTVT 675
Qy 661 YDLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQPCIEDPSPVILRLNF 720
Db 676 FDLALDHGRSLTRALFKETKTRALTRVKTGLNKHCSVKULLPACVEDSVPTITLRLNF 735
Qy 721 SLVGTPLSAFGLRVLAEADQRLFTALPFEKKNCGNDNICDDLSITFSFMSLDCLVVG 780
Db 736 SLVGVPISSQLNLQPLAVDDQTYFTASLPFEKKNCGADHICODDLSVVFPGPDLKTLVG 795
Qy 781 GPRENVTVTVRNDGEDSVRTQVTPFFPDLDSYRKVSTLQ-----NORSQRWR 829
Db 796 SDLELVNTVTVSNDGEDSVGTVTLYFPVGLSFRFAEGQVFLRKEDQQRGQSHLH 855
Qy 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLWSTSCSRHVFIRGGSQMTFLVTFDVSFPAELGDRLLRLRVGSE 913
Qy 890 NMMPRTNTEFOLELPVKYAVVMVTSHGVSFTKYNLFTASE-NTSRVNMHQYQVSNLQOR 948
Db 914 NNVPCTKTTOLELPVKYAVYTMISSHDQFTKYNLFTSEKETSUVVEHFRQVNNLQOR 973
Qy 949 SLPISLFLVPVRLNQTVMWRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCVS 1008
Db 974 DVPVSINFWPIELKGEAVW-TVMVSHFQNPPLTQCYNRLKPTQPDLLTHMQKSPVLDCS 1032
Qy 1009 TAVCQRIODIPFFGIOREFNATLKNLSFDWYIKTSHNHLLIVSTAILENDSVFTLLP 1068
Db 1033 TADCLHLRCDIPSLGILDELVPILKGNISFGMISQTLQKVKLLSEAEITNTSVYSQLP 1092
Qy 1069 GOGAFVRSOTETKVPFVFPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMNSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKHNVPVPLIVGSSVGGLLLLALITAILYKAGFFKRYQKEMLEE 1152

RESULT 6
Q9QYE7
ID Q9QYE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QYE7,
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSP; F11215; IBHQ.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 12660 MW; 2258491A984A705E CRC64;

Query Match 55.2%; Score 3244.5; DB 11; Length 1161;
Best Local Similarity 57.6%; Pred. No. 5.7e-232;
Matches 649; Conservative 163; Mismatches 302; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFOENARGFGQSVVQLQSRVVVGAPQEIVAANQRGSLYQCDYSTGSCBPIR 61
Db 21 NLDVEEPIVREDAAAFQIVVQGGSLVVGAPLEAVAVNQTGLDCAPATGMCQPIV 80
Qy 62 LQVPEAVNMSLGLSLAATSPOLLACGPTVHTQTSNTYVVKGLCFGLFNSLRQQPKF 121
Db 81 LRSPLAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNYAKGSCLLGSSL-QFIQAV 139
Qy 122 PEALRGCPQSDSDIAFLVDSGSIIPHDFRAKEFI1TVMEQLKSKTLESLMOYSEFR 181
Db 140 PASMPECPROEMDIAFLIDSGSGINQRDFAQMDQFVKALMGFEFASTTLESLMOYSILK 199
Qy 182 IHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVVRELFTNGARKNAFKILIT 241
Db 200 THFTTFEKNILDPQSLVDPIVLQGLTYTATGIRTVMEELFHSKNGSRKAKILLVIT 259
Qy 242 DGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELMTVAKPPRDHVQINN 301
Db 260 DGQKYRDPLEYSVDVIPAADKAGIIRYAGVDAFQEPALKELTIGTSAPPQDHVFKVGN 319
Qy 302 FEALKTIONQLREKI FAIEGTOTGSSSSFEHMSOEGFSAITNSGFLLTGVSQYDAG 361
Db 320 FAALRSIQRLQEKIFAIEGTQSRSSSFQHEMSQEGFSSALTSQDGLVAVGVSFWSG 379
Qy 362 VFLYTSKEKSTFIMNTVDSMDNDAYLGAAAAIILNRVOSLVGAPRYQHIGLVAMFRQ 421
Db 380 AFLYPNTRPTFINMSQENEDMRDYLGYSTALAFWKGVHSLILGAPRHQHTGKVIPTQ 439
Qy 422 NTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQVSCPLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSGLVLIGAPHYYEQTGGQVSCPLP 499
Qy 482 RQBARWOCDAVLGEQGPWGRFGAALTVLGVNGDKLTVAIGAPGEENKGAVYLF 541
Db 500 -GVGRWQCEATLHGEQHPWGRFGAALTVLGVNGDNLADVAIGAPGEESRGAIVIFP 558
Qy 542 GTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 601
Db 559 GASRLIMPSPQRISASQIPSRIQYFGQSLSGGQDLTMDGLVDLAVGSKRVLRLSLP 618
Qy 602 VLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVV 661
Db 619 LKVELSIRAPMEVAKAVVQCWERTPTVLEAGATVCLTVHKGSPDLL--GNVQSVRY 676
Qy 662 DLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQPCIEDPSPVILRLNFS 721
Db 677 DLALDPGRLLSRAIFDETCKNCTLGRKTLGLGDHCEVTKULLPDCVEDAVSPIILRLNFS 736
```

```

QY 722 LVGTPLSAFNGNRPVLADDAQRLFTALPFPEKNGCNDNICODDLSITSPMSLDCLVVG 781
Db 737 LVKDSASP-RNLHPVLAVGSDHIITASLPFEKNCQKELLCEGLGIFSNGLQVLVVG 795
QY 782 PREPNVTYVRNDGSDSYRTQVTFPEPLDLSYRKVSTLQNSORSORSLACESASSTVS 841
Db 796 SPELTVTVVMNEGSDSYGLVKFYFAGLSYRRTGTQ-QHQPVLRLACEAPAAQED 854
QY 842 GALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTBFQ 901
Db 855 --LRSSCSINHPIIFREGAKTTFMIFDVSYKAFGLDRLLRKAKASNNKPDNTKTAQ 912
QY 902 LELPKYAVYVMTSHGVSTKYLNFTASENTSR-VWQHQQVSNLQORSPLISLVLPV 960
Db 913 LELPKYTYVTLISRQSDTNHFNSSHGGRROEAAHRYRNNLSPLKLAVRNFWFV 972
QY 961 RLNQTVINDRPOVTSSENLS--TCHTKERLPSHDSFLAELRKAPVNVCSIAVCQRIQCDI 1019
Db 973 LLNGVAVMD---VTLSSPAQGVSCVSKMPQNDPDLTOIQRSLVDCSIADCLHFRCDI 1029
QY 1020 PFFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db 1030 PSLDIODELDFILRGNLSFGWSQTLQEKVLLVSEAITFTDTSVYSQLPQGEAFURAQVE 1089
QY 1080 TKVEPEVNPPLVGVSSVGGLLLALITAAALYKLGFFKROYKDM 1126
Db 1090 TTEELYVVEPIFLVAGSVGGLLLALITVVLVYKLGFFKROYKEML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TrEMBLrel. 12; Created)
DT 01-NOV-1999 (TrEMBLrel. 12; Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21; Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.1%; Score 1534.5; DB 11; Length 1161;
Best Local Similarity 34.0%; Pred. No. 9.6e-105;
Matches 399; Conservative 217; Mismatches 456; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFOENA-RGFQCSVVVQLGSRVAVGAQPEIIVAANQRGSLYQCDYSTGSCPE 59
Db 24 YNLDRTPQSLAQAGRHFGYQLQIEDG-VVVGAPCE---GDNTGLYHCRITSBFCQP 79
QY 60 IRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCLFLGSLNLRQOPQ 119

```

```

Db 80 VSLH-GSNHTSKYLGWMTLATDAAGSLACDGLSRTCDQNTVLSGLCYLFPQSLEGPML 138
QY 120 KFPBALRGCOESDIAFLVDGSGSIIPHDFRAKEFISTVMEOLKSKTLFSLMQYSEE 179
Db 139 QNRPAYOECKMGKVDLVLEFDGSGSLDRKQKOFKILEFMKQVMRKLSNTSYQFAAVQSTD 198
QY 180 FRIHFTPEF-QNNPNPRLSKIPITOLLGRTHATGIRKVVVRELNFINTGARKNAFKILI 238
Db 199 CRTEFTEDVYKQKNPDVLLGVSQPMFLTNTFRALNVVAVHFKESGARPDATKVLV 258
QY 239 LITDGEFGDPLGYEDVIEADREG-----VIRYVIGVDGAFRSKSKQELNVTASKP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIGICKHFVSQKOKTLHFASEP 304
QY 292 PRDIVQINNFEALKTTQNLQREKIFAIEGTQTSSTSSPHEMSQEQFSAAITNGPLLS 351
Db 305 VEEFVKLIDTPEKLDLQDLQRIYAIEGTNRODLTSFNMELSSSSISADLSKGHAVWG 364
QY 352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGYAAA-IILNRNVQSLVLGAPR 409
Db 365 AVGAKDWAGGLDREDLQATFVQEPDLSVDRGGYLGTVVAMWTSRSSRPLLAAGAPR 424
QY 410 YQHIGLVAMER--QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHY 467
Db 425 YQHVQVLLFOAPEAGGRWNTQKIETQIGSYEGGELCSVDLDQDGEAELLIGAPLFF 484
QY 468 BQTRGGQVVCPLPRGORARWQCDVLYGEOGQWGRFGAALTVLGDVNGDKLTDVAIGA 527
Db 485 GEORGRGVFTY---QRRQSLFEMVSELQGDGPGYPLGRFGAAITALTIDINGDRLTDVAVGA 541
QY 528 PGEEDNGAVYLFHGTSGSISPSHSORISAGSKSLSPRLQYFGOSLSGGODLTMDGLVDLT 587
Db 542 PLEE--QGANVIFNGKPG-GUSPQPSQRIQQAQVFPGRWFGRIHGVKDLGGDELADV 598
QY 588 VQAQHVLLLRSPQLRVKAIEMFNPREVARNVFECNDQVQVKGKAG-EVRVCLHVQKST 646
Db 599 VGPEGRVVVLSRRPVDVVTLSFSPEIIPVHEVECSYSAREEQHKVKKACFRKPLT 658
QY 647 RDRUREGQISVVYDLDALDSGRPHSRVAFNETKNSRTRQTVGLGLTCTETLKLQLPNC 706
Db 659 PQ--FQGRLLANLSYTLQLDGHRMRSGLFPDGGSHELSGNTSITP-DRSCLDLHFHFPIC 715
QY 707 IEDPVSPIVLRNLSLV---GTPLSAGCN-LRPVLAEDAQRLFTALPFPEKNCNDNLCO 762
Db 716 IODLISPIVNSLFLSLEEETPRQKQKQAPILRPSIHV-TKEIPFKNCGEDKKCE 774
QY 763 DDLSTIFSMSLDCLVWGGP-----REFNVTVTVRNDGSDSYRTQVTFPFLDLSYRKV 816
Db 775 ANLTLSPPARS-----GPLRLMSSASLAVEMWTLNSGEDAYWVRLDLPRLGSLFRKV 827
QY 817 STLQNRORSWRRLACESASSTEVSGAL-KSTSCSINHPIIPENSEVFNITFDVDSKAS 875
Db 828 EMLQ--PHSRMPVSCEEL--TEGSSLLTKLTKCNVSSPIFKAGQEVSLQVMFNTLLSS 882
QY 876 LGNKLKLLKANVTSEN-NMPTNKTFFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSR 934
Db 883 WEDFVELNGTVHCENENSSLOQEDNSAATHIPLYVLPNLTKEQENSTLYISTPKGPKTQ 942
QY 935 VMQHQQV---SNLQORSPLISLVLPVRLNQTVIWRPQ--VTFSENLS--TCHT 985
Db 943 QVQHVVYQVRIQPSAYDHNMP-TLEALVGPV-----WPHSEDPITYTWSVQTDPLVTCHS 995
QY 986 KE-RLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFQIGIEFNATLKGNSLFDWYIKT 1044
Db 996 EDLKRPSSE--AEQPCPLPGV-----QFRCPIVF---RREILIQVTGTVELSKEIKA 1041
QY 1045 SHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTEKVEFPEVNPPLVLCVSSVGGLL 1104
Db 1042 S-STLSLCSLSVSNSSKHFHYLGSKA-SEAQVLVKVDLIHEKMLHVVLVSGGLVL 1099
QY 1105 LALITAAALYKLGFFKROYKDM-SEGGPPGAEP 1136

```


DR InterPro: IPR002035; VWF_A.
DR Pfam: PF001839; FG-GAP; 5.
DR Pfam: PF000357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50234; VWF; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 24.0%; Score 1409; DB 13; Length 1196;
Best Local Similarity 31.5%; Pred. No. 2.2e-95;
Matches 375; Conservative 224; Mismatches 447; Indels 144; Gaps 39;

QY 1 FNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQEIIVAAHQSLYQCDYSTGS 56
DB FNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQEIIVAAHQSLYQCDYSTGS 56
DB FNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQEIIVAAHQSLYQCDYSTGS 56
QY 57 CEPRLQVP----VEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVGKLCPLFGS 112
DB CEPRLQVP----VEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVGKLCPLFGS 112
DB CEPRLQVP----VEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVGKLCPLFGS 112
QY 89 CQ--RLQPGSESVRFEGMSAAVSSAALTS-----CSPVFPHECDGNSYLVGVCYQFSS 140
DB CQ--RLQPGSESVRFEGMSAAVSSAALTS-----CSPVFPHECDGNSYLVGVCYQFSS 140
DB CQ--RLQPGSESVRFEGMSAAVSSAALTS-----CSPVFPHECDGNSYLVGVCYQFSS 140
QY 113 NLROQKQKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFIIVMEQLKSKTLES 172
DB NLROQKQKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFIIVMEQLKSKTLES 172
DB NLROQKQKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFIIVMEQLKSKTLES 172
QY 141 SL-QAVSNFTAAQECSEKREVLNVLFLDGGSSMKTYDFEMKNKFIKIMKLSNSIKFA 199
DB SL-QAVSNFTAAQECSEKREVLNVLFLDGGSSMKTYDFEMKNKFIKIMKLSNSIKFA 199
DB SL-QAVSNFTAAQECSEKREVLNVLFLDGGSSMKTYDFEMKNKFIKIMKLSNSIKFA 199
QY 173 LMOYSEEFRIHFTFKFQNNPNRSLIKPITOLLGRTHATGRKVVRELFN-ITNGARK 231
DB LMOYSEEFRIHFTFKFQNNPNRSLIKPITOLLGRTHATGRKVVRELFN-ITNGARK 231
DB LMOYSEEFRIHFTFKFQNNPNRSLIKPITOLLGRTHATGRKVVRELFN-ITNGARK 231
QY 200 AVQFSDTRVTFDFNDQSGSAEKLMBE-TMKSLTNTHKALDYILKNLLSMLSGADS 258
DB AVQFSDTRVTFDFNDQSGSAEKLMBE-TMKSLTNTHKALDYILKNLLSMLSGADS 258
DB AVQFSDTRVTFDFNDQSGSAEKLMBE-TMKSLTNTHKALDYILKNLLSMLSGADS 258
QY 232 NAFKILILITDGEKFGDPLGYED--VIPEADREGVIRYVIGVDAPFSEKSRQELNTVAS 289
DB NAFKILILITDGEKFGDPLGYED--VIPEADREGVIRYVIGVDAPFSEKSRQELNTVAS 289
DB NAFKILILITDGEKFGDPLGYED--VIPEADREGVIRYVIGVDAPFSEKSRQELNTVAS 289
QY 259 KAQKALVITD---GQPSDNDYVNLKKCDEQNILRYIIGV-----KVDLITELQLAS 309
DB KAQKALVITD---GQPSDNDYVNLKKCDEQNILRYIIGV-----KVDLITELQLAS 309
DB KAQKALVITD---GQPSDNDYVNLKKCDEQNILRYIIGV-----KVDLITELQLAS 309
QY 290 KPRDRHVQINFEALKTIONQREKIFAIEGTQTGSSSFEHEMSQEGFSAITNGPL 349
DB KPRDRHVQINFEALKTIONQREKIFAIEGTQTGSSSFEHEMSQEGFSAITNGPL 349
DB KPRDRHVQINFEALKTIONQREKIFAIEGTQTGSSSFEHEMSQEGFSAITNGPL 349
QY 310 EPRKNTFYIKDYSKGLGLDNLQKKIYNEGSDVDAQDRQKELSQSGSVVYQESVI 369
DB EPRKNTFYIKDYSKGLGLDNLQKKIYNEGSDVDAQDRQKELSQSGSVVYQESVI 369
DB EPRKNTFYIKDYSKGLGLDNLQKKIYNEGSDVDAQDRQKELSQSGSVVYQESVI 369
QY 350 LSTVGSYDAGGVFLYTSKEKSTFINNRVDSMN-DAYLGAAAIILNRVQSLVLGAP 408
DB LSTVGSYDAGGVFLYTSKEKSTFINNRVDSMN-DAYLGAAAIILNRVQSLVLGAP 408
DB LSTVGSYDAGGVFLYTSKEKSTFINNRVDSMN-DAYLGAAAIILNRVQSLVLGAP 408
QY 370 VGSVGNDRGALYEVGT-SGSDFKETEIIDPAVKDSYNGYSTVLGMRGVSLLSGAP 428
DB VGSVGNDRGALYEVGT-SGSDFKETEIIDPAVKDSYNGYSTVLGMRGVSLLSGAP 428
DB VGSVGNDRGALYEVGT-SGSDFKETEIIDPAVKDSYNGYSTVLGMRGVSLLSGAP 428
QY 409 RYQHIGLVAFRQNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYE 468
DB RYQHIGLVAFRQNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYE 468
DB RYQHIGLVAFRQNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYE 468
QY 429 RAEHTGLVTLFTKQNTWTVTSTNINEGQIGSYFGASLSDLDVSDGSDPFLVGLPFLYQ 488
DB RAEHTGLVTLFTKQNTWTVTSTNINEGQIGSYFGASLSDLDVSDGSDPFLVGLPFLYQ 488
DB RAEHTGLVTLFTKQNTWTVTSTNINEGQIGSYFGASLSDLDVSDGSDPFLVGLPFLYQ 488
QY 469 QTRGGQSVCLPRGQARQWOCDAVLYGEO-----GQWGRFGAALTVLGDVNGDKLTV 523
DB QTRGGQSVCLPRGQARQWOCDAVLYGEO-----GQWGRFGAALTVLGDVNGDKLTV 523
DB QTRGGQSVCLPRGQARQWOCDAVLYGEO-----GQWGRFGAALTVLGDVNGDKLTV 523
QY 489 SQ-----PRTEGRLYVYSLSEQKYFQKTLNVQSSTTGRFAASVSLKDLNGDGLSDV 540
DB SQ-----PRTEGRLYVYSLSEQKYFQKTLNVQSSTTGRFAASVSLKDLNGDGLSDV 540
DB SQ-----PRTEGRLYVYSLSEQKYFQKTLNVQSSTTGRFAASVSLKDLNGDGLSDV 540
QY 524 AIGAPGEDNRGAVYLPFGTSGSGISPSHS-ORTAGSKLSPRLQYFGQSLSGQDGLTMDG 582
DB AIGAPGEDNRGAVYLPFGTSGSGISPSHS-ORTAGSKLSPRLQYFGQSLSGQDGLTMDG 582
DB AIGAPGEDNRGAVYLPFGTSGSGISPSHS-ORTAGSKLSPRLQYFGQSLSGQDGLTMDG 582
QY 541 AVGAPLE--NEGVIYIYLGORTGINPELTPQRLSVQSVLPGLQOFQSVLGTQMDMNDN 598
DB AVGAPLE--NEGVIYIYLGORTGINPELTPQRLSVQSVLPGLQOFQSVLGTQMDMNDN 598
DB AVGAPLE--NEGVIYIYLGORTGINPELTPQRLSVQSVLPGLQOFQSVLGTQMDMNDN 598
QY 583 LVDLTVGAQGHVLLRLSQPLRVKAIIMEFNPREVARNVFCNDQWVGKEAGEVVRCLHV 642
DB LVDLTVGAQGHVLLRLSQPLRVKAIIMEFNPREVARNVFCNDQWVGKEAGEVVRCLHV 642
DB LVDLTVGAQGHVLLRLSQPLRVKAIIMEFNPREVARNVFCNDQWVGKEAGEVVRCLHV 642
QY 599 LTDIVGAQGVILLKARPVMSVSAQLSFSPKEISLNYFECPGS--NAFNAFLNLTSCFTV 656
DB LTDIVGAQGVILLKARPVMSVSAQLSFSPKEISLNYFECPGS--NAFNAFLNLTSCFTV 656
DB LTDIVGAQGVILLKARPVMSVSAQLSFSPKEISLNYFECPGS--NAFNAFLNLTSCFTV 656
QY 643 QKSTRDLREGQITQSV--VTVYDLALDSGRPHSRVAFNETKNSTR--RQTOVLGLTQTCET 698
DB QKSTRDLREGQITQSV--VTVYDLALDSGRPHSRVAFNETKNSTR--RQTOVLGLTQTCET 698
DB QKSTRDLREGQITQSV--VTVYDLALDSGRPHSRVAFNETKNSTR--RQTOVLGLTQTCET 698
QY 657 TERTSS---TGSLEKLNVLNLDVVRGMSRGFFDQSSVSSRTLQOQSVLLDSGSCFN 713
DB TERTSS---TGSLEKLNVLNLDVVRGMSRGFFDQSSVSSRTLQOQSVLLDSGSCFN 713
DB TERTSS---TGSLEKLNVLNLDVVRGMSRGFFDQSSVSSRTLQOQSVLLDSGSCFN 713
QY 699 LKQLPNCIEDPVPIVLRNLFSLVGTPLSAFGLNLRVLAEDAORLFTALFPPEKNCND 758
DB LKQLPNCIEDPVPIVLRNLFSLVGTPLSAFGLNLRVLAEDAORLFTALFPPEKNCND 758
DB LKQLPNCIEDPVPIVLRNLFSLVGTPLSAFGLNLRVLAEDAORLFTALFPPEKNCND 758
QY 714 FSIFMLRCVADTVSFLKIRMFNS--QTEMLS--GNSAVLDVHSSTEENVEVFORNCNSN 770
DB FSIFMLRCVADTVSFLKIRMFNS--QTEMLS--GNSAVLDVHSSTEENVEVFORNCNSN 770
DB FSIFMLRCVADTVSFLKIRMFNS--QTEMLS--GNSAVLDVHSSTEENVEVFORNCNSN 770
QY 759 NICODDLISITFSFMSLCLVGGPREFNVTIVRNDGEDSYRTQVTFPFLDLISYRKVST 818
DB NICODDLISITFSFMSLCLVGGPREFNVTIVRNDGEDSYRTQVTFPFLDLISYRKVST 818
DB NICODDLISITFSFMSLCLVGGPREFNVTIVRNDGEDSYRTQVTFPFLDLISYRKVST 818
QY 771 NSCVADLKLNFSTFN--NTLVVENQAHFTVQVSLANPGDSDSYNTSIVLHYPEGISLSKFDA 829
DB NSCVADLKLNFSTFN--NTLVVENQAHFTVQVSLANPGDSDSYNTSIVLHYPEGISLSKFDA 829
DB NSCVADLKLNFSTFN--NTLVVENQAHFTVQVSLANPGDSDSYNTSIVLHYPEGISLSKFDA 829
QY 819 LQNRQSRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDV---DSKAS 875
DB LQNRQSRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDV---DSKAS 875
DB LQNRQSRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDV---DSKAS 875

DB 830 IKPSRTR-----SSCGDRDSCATNRRTTCSIDLPPVRSCTTTQFLGTGTPRWMKWDNDSS 881
QY 876 LGNKLKLLKANYTSNNMPTNKTEFOLELPVKYAVYVWVTSBGV-STKYLNFTASENSTR 934
DB 882 NRMEITMITANDNGNM---SDTEVRRSPVQVFAVDLAISLVAEDSVTYMNFSLSEDGPK 938
QY 935 VMOHQYQVSNLQORSLPISLFLVLPVR-----LN 963
DB 939 PLNITYKVENSGRKLDPVSVTLTLPQTPHVLTPHTFSMEVHVSFISSVHOIMCLLN 998
QY 964 QTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQ-----C 1017
DB 999 KHLFFSPELSAAVQVVRSTGWSLREVC-SQF--DLNKSAAVHFNLTADARLQNVKEYES 1055
QY 1018 DIPFFGIQEE--FNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGGAPVR 1075
DB 1056 KYSPYEFKDNVFSIAELNYSNTSLYNQTS-----SELKYNPH-----R 1094
QY 1076 SQTETKVPPEFVNPPLIV-GSSVGGLLLLALITAALYKLGFFPKROYKD 1124
DB 1095 SQTETKVPPEFVNPPLIV-GSSVGGLLLLALITAALYKLGFFPKROYKD 1124
DB SQTETKVPPEFVNPPLIV-GSSVGGLLLLALITAALYKLGFFPKROYKD 1124

RESULT 10
Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AAH08777.1; -
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRINA.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS50234; VWF; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6FF2546E8C632F9 CRC64;

Query Match 23.1%; Score 1359.5; DB 4; Length 1086;
Best Local Similarity 32.3%; Pred. No. 9e-92;
Matches 378; Conservative 185; Mismatches 431; Indels 175; Gaps 37;

QY 1 FNLDTENAMTFQ--ENARGFQSVVQLOGSRVVGAPQEIIVAAHQSLYQCDYSTGSCS 58
DB 26 YNLDVGRGARSFPGRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGCHL 81
QY 59 PIRLQVPVBAVNMVSLGLSLAATSPOLLACGPTVHQTCSNTYVVGKLCPLFGSNLRQOP 118
DB 82 PVTLR-GSNVTSKYLGMTLA--TDP----- 103
QY 119 QKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFIIVMEQLKSKTLESFSLMOYSE 178
DB 104 -----TDGS----- 117
QY 179 EPRHFTFEFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKVAFKILI 238
DB ----- 238

Db 118 SYKTEPFDSDYVKKRDPDALLKHVKHMLLLTNTFTGAINVATEVFEELGARPDAKVLII 177
Qy 239 LITDGEFGDPLGVEDVPEADREGVIRYVIGVGDAPFSEKSEKQELNTVASKPRDHVFO 298
Db 178 IITDGE--ATDSGNIDAOK-----IIRYIIGIKGHFQTKESQETLHKFASKASBFVKI 230
Qy 299 INNEFALKTIQNLREKIFAIEGTQTGSSSFEHEMSEQEGFAAITSNGPLLSVGSYDW 358
Db 231 LDITPEKLDLFTLEKQKIVIEGTSKQDLTSFNWELSSGSIADLSRGAHVAVGAKOW 290
Qy 359 AGGVF-LYTSKEKSTFFINTRVSDMDNDAYLGAAA-IILNRVQSVLVCAPRYQHIGLV 416
Db 291 AGGFLDLKADLQDDTFIGNEPLTPEVRAGLVGVTVWLPSPRQKSTLLASGAPRYQHMGV 350
Qy 417 AMFR-QNTGWESNANVKGTGATGAYEGASICSVDVDSNGSTDVLVLGAGHYEQRGGQ 474
Db 351 LLFQEPQGGHWSQVTHGTQSGYFGBELCGVDVDDGETELLGILGAPLFGEORGR 410
Qy 475 VSVCLPRGRARWQCDV--LYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEED 532
Db 411 VFIV-----QRQLGFEVSELOQDCPYPLGRFGEAITALTDINGDLVDVAVGAPLEE- 464
Qy 533 NRGAVYLFHTSGSGISPSHSQRTAGSKLSPRIQYFGOSLGGODLTMDGLVLTGVAQG 592
Db 465 -QGAIVYIFNGRHG-GLSPQSPQRTEGTQVLSGIQWFGRSIHGVKDLGDLGADVAVGAS 522
Qy 593 HVLLRSQPLRVKAINFENPREVARNVFECNDVV-KGEAGEVRVCLHVQKSTRDLR 651
Db 523 QMIVLSRPVDMVTLMSPFAETIPVHEVECSYSTSNMKEGVNITICQI-KSLIPQF- 580
Qy 652 EGOISVVTYDLDLDSGRPHSAVFNETKSTROTGLTGTOTCETLKLQPNCTEDPV 711
Db 581 QGRUVANLTTLQDGHTRRRGLFPGRHELNRNIAT-TSMSCDTFSFHPVVCVDLI 639
Qy 712 SPIVLRNLSL--VGTPLSAFGN-----LRPVLAEADAORLEFALPPEKNCNQNICQ 762
Db 640 SPINVSLSFWESEGTPRORAGKDIPILRPLSHSETWEI-----PFEKNCEDKCE 694
Qy 763 DDLSTITSFMSLDCLVGGPRFNVTVVRNDEGDSYRTQVTFPPDLDLRYKVKSTLQNG 822
Db 695 ANLRVSFSPARSRALRLTAFASLSVELSLNLEEDAYWQDLHFPPLGSLFRKVEML-- 751
Qy 823 RSQSRWELACES--ASSTEVSGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKL 880
Db 752 KPHSQIPVSCDELPEERLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNWSGDSV 808
Qy 881 LLKANVTSENN---MPTKNTFQLEPLVKYAVMVVTSHGVSSTKYLNFTASENTSVM 936
Db 809 ELHANVTNNEDSDLEDNSATTI--IPILYPINILIQOEDSTLVSTPKGPKHQV 865
Qy 937 QHOYQV---SNLQORSIP-LSLFLVPVRLNQTVWRDPQVTFSENLSSTCHTK--ERLP 990
Db 866 KMYQVRIQPSIHDNIPITLEAVVGVPPPEGPITHQWSVQMEPPV--PCHVEDLERLP 923
Qy 991 SHSD--FLAELRKAPOVNCIAVCORIQDIPFGIGEEFNATLKNLSFDWIKTSHNH 1048
Db 924 DAAPCLUPGALFCPPV-----FRQETLVQVIGTLELVGEIAS-SM 964
Qy 1049 LLIVSTAEILFNDSVFTLLPQGAFAVRSQETKVEPEVENPPLIIVGSSVGGILLILI 1108
Db 965 FSLCSSLISIFNSKSHFLYGSNASL-AQVMKVDVYVYKQMLYLYVLSGIGILLILLI 1023
Qy 1109 TAALYKLGFTFKROYKMMSEG-GPPGAEP 1136
Db 1024 FIVLYKVGFFKRNLEKNEAGRGVPGIP 1052

RESULT 11

Q98TF0

ID

AC

Q98TF0;

DT

01-JUN-2001

(TReMBLrel. 17, Created)

DT

01-JUN-2001

(TReMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DB CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1; .
DR HSSP; P20701; ILPA.
DR InterPro; IPR001969; Aspartate site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00357; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 23.1%; Score 1358.5; DB 13; Length 1187;
Best Local Similarity 31.6%; Pred. No. 1.2e-91;
Matches 377; Conservative 206; Mismatches 452; Indels 157; Gaps 43;

Qy 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQRGSLYQCDYSTGS 56
Db 32 FNIDTEHPLRFNGTPTDFFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSTADLQS 88
Qy 57 CEPRLQVP-----VEAVNMSLGLSLAATTSPOLLACGTVTHTGTCENTYVKGCLFLFGS 112
Db 89 CK--RLQRFGESESVRFEGMSAAVSSAALTS-----CSPYFAHECDGNSYLVNGCYQFNS 140
Qy 113 NLRQOPKQPEALRGCPQSDSDIAFLVDGSGSIIPHDFFRAKEFIITVMBOLKSKTLFS 172
Db 141 SL-QAVSNFTAAQECSEKREVNLFVLDGSSSKWAVEFDNKNKFIKDVMMKLSNSIKFA 199
Qy 173 LMQYSEEPRIHTFKEFQNNPNRSLIKPTQLLGRTHATGIRKVVRELFN-ITNGARK 231
Db 200 AVQFSTEIRTVDFDNDYQNGSABEKLME-RHMKSLTNTYKAINVYLKNVLSVSGADP 258
Qy 232 NAPKILILITDGEKFGDPLGYED--VIPADREGVIRYVIGVGDAPFSEKSRQELNTVAS 289
Db 259 NAQKALVITD-----GDPDNDNDYINILNICDEQNILRYIIGV-----KVDLTTLTQLAA 309
Qy 290 KPRDHVFOINNFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSEQEGFAAITSNGPL 349
Db 310 EPKLNNTFYIQEYNGUKLGLDNLQKXIYNIEGSEKAHGRDRQKELSGSGSVVYQESVI 369
Qy 350 LSTVGSYDWAGGVFLYT---SKEKSTFINNTRVDSMDNDAYLGAAAIIILNRVQSVLVLG 406
Db 370 VGSVGSNDRGALYEWGSGSKFRQTEITDPVN---KDSYMGYSTVLGRHGVSLFLSG 426
Qy 407 APRYQHILGVAMPFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHY 466
Db 427 APRAHTGLVTLFTKNESTWTVMRNINEGQISYFGASLSDLDVDSGDSDFLLVGAFLF 486
Qy 467 YE-QTRG-QGVSVCLPLPRQARWQCDVLYGEQ--GQPMGRFGAALTVLGDVNGDKLTD 522
Db 487 YSQPRAEGRLYVYTL-----SEYQSKTLQSTTGRFATSLASLKLQNLGDLSD 535
Qy 523 VAIGAPGEEDNRGAVYLFHGTSGSGISPSHS-ORLAGSKLSRLQYFGOSLGGQDLTMD 581
Db 536 VAVGAPLE--NEGVTYIYLDGTHGINPEHAPORIPARVLPGLQOQFGVSLSGQMDND 593


```

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match          19.9%   Score 1167.5; DB 11; Length 1167;
Best Local Similarity 29.0%   Pred. NO. 2e-77;
Matches 358; Conservative 214; Mismatches 456; Indels 205; Gaps 43;

QY 1 ENLDTEA--MTFQENARFGQSVVLOGSRVVVGCAPQEIIVAAANORG-----LYQCDY 52
DB 20 FNMVDVMAWVATALQGPVAVLSLLHLDPSN-----NOTCLLVARRSNRNTAALYRCAI 74
QY 53 STGSCPTIRLOVPEAVNMSGLSLAATT--SPOLLAC-GPTVHQTCSNTYKGLCLF 109
DB 75 SI-SPELACQ-PVEHICMPKRGYQVTLVGNHNGVLVCIQVQARKFSLNSELTGACSL 132
QY 110 FGSNLRQOPQKPEALRG-----C-----POE 131
DB 133 LTPNLDLQAQVYFSDLEGFLDPGAHVDSGDYCRSGKSTGBEKKSAARRRTVEEDED 192
QY 132 DSDIAFLVDGSGSIIPHFRAKERISIVMQL--KSKTFLSLMOYSEEFRIHPTKPEF 189
DB 193 GTEIAVLVDGSGSIEPSDFQAKNFISTMRNFYEKCEFCNPFALVQYGAIVQTEFDLQES 252
QY 190 ONNPNRSLPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILILITDGEKFGDP 249
DB 253 RDINASLAKVOSIVQKEVKTASAMQHVLDNIFIPSRGSRKALKVMVLTGDIFGDP 312
QY 250 LGYEDVPEADREGVIRVIGVDFAFRSEKSRQELNTVASKPRDRHVFQINNFEALKTQ 309
DB 313 LNLTVINSKPMQVVRGAIGVDFAKNNVTRELKLIASDPKEAHTFKVTNYSALDGLL 372
QY 310 NOLREKIFAIEGTQGTSSSEHEHMSQEGFSAITSNGP-LLSTVGSVDWAGGVFLY-TS 367
DB 373 SKLQORIIVHMEGT---VGDALQYLAQTGFSAQILDKQVLGTGVAFNWSGGALLYSTQ 429
QY 368 KEKSTFIWNT-RVDS-DNNDVGLVAAAIILNRNVQSLVGLAPRYOHIGLVAMPFRONTGM 425
DB 430 NGRGCFNLQAKEDSRTVOYSYLGSLVLAHLKHAIGISVAGAPRHLKRGAVFELKREDR- 488
QY 426 WESNA---NVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVSCPLPR 482
DB 489 -EEDAFVRIIEGEQNGSVFGSVLCPVDIDMDGTDFTDLVAAPFYHIRGEGRVYVQYPE 547
QY 483 QORARWQCDVLYGSGQGPWRGFGAALTVLGVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLSAHTLSGHPGTLNTRFGFMAAAGVDINODKFTDVAIGAPLEGFGAGDGSYGS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLL 596

```

```

DB 607 VYIYNHSG--GLYDPSQQRASSVAGSLHYFGMSVGGDLDFNGDGLADITVGRSDSAV 665
QY 597 LRSQPLVRVKAINEFNPREFARNVFECDQVVKGEAGEVRVCLHVQKS----TDRRLREG 653
DB 666 LRSRPVVDLTVSMFTFP-----DALPMVFIGKM--DVLNCFEVDSSVVASPFGLREM 715
QY 654 QIQSWVTVYDIALDSGRPHSRAVFNETKNSTRRTQVLGLTQTC-----696
DB 716 FLNFTVDV-----TKQRLQCEDESGGSCCLRKWNGSFLCEHPWLI 760
QY 697 ETLKQLPNCIEDPSPVILRLNFSLVGTPLSAFENLR-----PVLAEADAQLFTALF--P 750
DB 761 STEEL-----CEEDCFSNITIKVYE-----POTSGRRDYENPTL--DHYKEPSAIFQLP 809
QY 751 FEKNCNDNICQDDLSTIFSPMSLDCLVVGSPREFNVTVVRNDGEDSYRTQVTFPPPLD 810
DB 810 YEKDKCKNVFCIAEIQLTN-ISQBELVVGVTKEVTWNISLTNSGDSYMTNMAALNPRN 868
QY 811 LSYRKVSTLQNRORSQSWRLACESASSTEVSGALKSTSCSINHPIFFENSEVTFTNITFDV 870
DB 869 LQFKKI-----QKPSVDPVQCDPKV---ASVLVNMCKIGHFIL-KRSSVNVSVTWQL 918
QY 871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVVMVYVTSHGVSXYLNFTASE 930
DB 919 EESVFPNRTADITVTISNSNEKSLARETR---SLQFRHAFIAVLSR--PSVMYNN--TSQ 971
QY 931 NTSRVMOHQVQSVNLGORSPLISLVFLVPVRLNQTIVWDRPQVTFSSNLST-----CHT 985
DB 972 SPSPDKHEFFFNHNGENLFGAVFQICVPKIQDF-----QIVRVKNLTKQDHTCTQ 1025
QY 986 KERLPSSHDLAEALKAPVNCSTAVCORIQCDIPFGIQEENATLKGNLSPDWYIKTS 1045
DB 1026 SQEPACSDPVQVHKEWHSVVCAL-----TSNKENTVVAEISVG 1065
QY 1046 HNHLIIVSTA-----EILFNDVSFTLLPGOGAFVRSQTEKVPFF-----EVPNPLPIV 1095
DB 1066 HTKQLLRDVSLELPILGEISFNKSLYEGINAE-----NHRKTIIVFLKEBETSLPLII 1119
QY 1096 GSSVGGGLLALITAAALYKLGFFKQYKDMSE 1128
DB 1120 GSSIGGLLVLLVITAILFCQGFRRKYQQLNLE 1152

RESULT 14
O88341
ID O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

```

DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;

Query Match
Best Local Similarity 19.1%; Score 1123; DB 11; Length 1167;
Matches 334; Conservative 207; Mismatches 421; Indels 244; Gaps 37;

QY 47 LYQDYSTGSCPIRLQVPVEAVNMSLGLSLAATSPOLLACGPTVH-----94
DB 67 LYQC-----ATSIIPDEIGCPVEHILMPKRYQGVY 98
QY 95 -----QTCSENTYVKGICLFGSNLRQOPKQFPEALRG-----127
DB 99 LVRNHNGLVLCIQVQSRKPRNSLSELTGACSLTLPNLDLQQAQYFSDLEGVLDLGASVNS 158
QY 128 -----CPQED-----SDIAFLVDGSGSIIPHDFRRAKEFI 157
DB 159 GDYSGKSGTGEETKSARLRQAQVEEEDBEAGTEIAI VLDGSGSIEPSDFQAKOFI 218
QY 158 STVMEQL--KSKTFLPSLMQYSEEFRIHFTFEFQNNPNRSLIKPIITOLLGRTHATGI 215
DB 219 STWNRNFEYKCEPCFNALVQYGVIGIOTEPDLDLSDRINASLAKVQSIQVQKEVTKTASAM 278
QY 216 RKVRELFINITGARKNAEKILLITDGEKFGDPLGYEDVIEADREGVIRVIGVDAP 275
DB 279 QHVDNIFIPSGSRKALKVMVLTGDI FRDPLNLTTVISSKMQGVVRAIGVGNAP 338
QY 276 REKSRQELINTVASKPRDHVQIINNFEALKITONOLREKIFAEIGTQTGSSSPHEHMS 335
DB 339 ENNVTYRELKLASDEPKAAHTFKVNYNSALDGLSLKLOQRIIHMEGT--VGDYLOQYOLA 395
QY 336 QEGFSAITNSGP-LLSTVSGVDWAGGVFLY-TSKEKSTFINNTRVDSMDNDA---YLGY 390
DB 396 QTGFSAIQILDKGQVLGTVGAFNWSGALLYNTQNGRGRFLNOT-AKEDFRAAQSYLGY 454
QY 391 AAAILLRNRVQSLVLCAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDV 450
DB 455 SVAALHKAHGVSVYAGAPRHKLURGAVFELOKEDGETFMRRIEGEQMGSYFGSLVCPVDI 514
QY 451 DNGSTDLVLIGAPHYEEOTRGQVSVCPPLPRQORARWQCDVLYEQCQOPWGRFGAALT 510
DB 515 NMDGITDFLLVAAPFTHIRGEGRVTVYRV-HEQDAPFSLVTLGYPLGTSRFGFAMA 573
QY 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGAVYLFPHGTSGSGISPSHSQRIAGSKLSPR 564
DB 574 AVGDINQDKFTDVAIGAPLEGFGAGDGAGSYGVYVYNGHSG-GLHASPSQOI RASSVALG 632
QY 565 LQYFGOSLSGGQDLTDGLVDLTVGAQHVLLRSOPVLVRKAIMFNPREVARNVFECH 624
DB 633 LYFYGMSVSGGLDFSGDDLADITVGSQDVAVLRSRPPVVDLTVMSTFTPDALP-----685
QY 625 DQVVKGEAGEVRVCLHVOKS---TRDLREGQIQSVVTVYDLDALDGRPHSRVAFNETKN 681
DB 686 ---MAPKDMVDELCKFVDUSSAVPSEPLRGHMSLNTFTVDVV-----TKQ 727
QY 682 STRRQ-----TQVLGLTQTCETLKLQPLNCIEDPVSPIVLRINFS 721
DB 728 KQRLOCADRSCCQCLMKWSSGSSLCHEHFLGISTEEEL-----CEDDCFSNITIKVSVE 781
QY 722 LVGTPLSAFGLNRPVLAEDAQRLFTALF--PFENKNGNDNICQDLSLITFSPLMSLCLVV 779
DB 782 F-QTSEERNHNPIL--DHYKEPSAIFQLPYEKCKNKVFCIAETIQLTAISQQD-LVV 837
QY 780 GGPREFNVTVVNDGEDSVRTQVTFPFLDLSYRKVSTLQNRQSRQSRWLACASASSTE 839
DB 838 GITKEVTMISLNSGEDSTMTNNALNYPNLOFKKI-----QKPLSPDIOCDQDKPV- 890
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKT 899

891 --ASVLVNMCKIGHPIIL-KRSSVNVSVTMQLESIFFPNRTADITVTISNSNEKSLARETH 947
900 FQLELPVKYAVYVMVTVTSHGVSTKYNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVP 959
948 ---SLOFRAHIAVLNR--PSVWYMN--TSOSSDHDKEFFFNHNGENHFGAVFQLOICVP 1000
960 VRLNQTVIWRPQVTFSENLSST-CHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCD 1018
1001 ITRDL-----QIRVKGHLTKQAH-----ECTQSQEPCTGSDPQVNV---1039
1019 IPFGIQOEFNATL-----KGNLSFDWYIKTSHNHLI--VSTAEIL-----FNDSVFTL 1066
1040 -----BEHSHVCAITSNKENVTVAABEISMGHTKQLLRDISELOILGEISFNKSLYEG 1092
1067 LPGAQAFVRSOTETKVEPF-----EVPNPLPLIVSGSVGLLALLALITAAALKGLGFFKQY 1122
1093 LNAE-----NHRKITYIFLKEEPHSLPLIIGSSIGGLLVLVIIAILFKCGFFKRY 1146
1123 KDMSE 1128
1147 KQLNLE 1152

RESULT 15
O42094
ID O42094 PRELIMINARY; PRT; 1171 AA.
AC O42094;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE ALPHA integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
RT the alpha1 integrin gene.";
RL J. Biol. Chem. 272:26643-26651(1997).
DR EMBL; AB000470; BAA23160.1; -;
DR EMBL; AB000471; BAA23161.1; -;
DR HSSP; P17301; IAOX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 5-
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Query Match
Best Local Similarity 17.9%; Score 1053; DB 13; Length 1171;
Matches 339; Conservative 216; Mismatches 513; Indels 144; Gaps 46;

QY 2 NLDTENAMTFQENARG-FQSQVVLQ---QGSVVVVGAPQEI VAAQNGSLYQCDYSTGSC 57
DB 19 NVDVKNMTFSGPLEDMFGYTVQQYENEEGKWLIGSLVPGQPEKRTGDKVYKPCVRDSQ 78
QY 58 EP-IRLQVP-----VEAV--NMSLGLSLAATSPOLLACGPTVHTQTCSENTYVKGILC 107
DB 79 SPCKINLNPDATSVPNVMEVKENMTLGTTL-VTPNKGGLACGLPYAYKCGRLHYTTGVC 137

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481b-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMWSEGGPGGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5862	99.8	1153	19 AAW65090	Human Beta-integrin
2	5862	99.8	1153	21 AAB07360	Human CD11b protei
3	5862	99.8	1153	23 ABG61469	Human Beta2 integr
4	5862	99.8	1153	23 AAU80252	Human integrin l a
5	5862	99.8	1153	23 AAO14428	Integrin Mac-1 alp
6	5852	99.6	1153	11 AAR04136	Alpha subunit of M
7	3460	58.9	1163	11 AAR07120	p150.95 alpha subu
8	3446	58.6	1163	19 AAW65091	Human Beta-integrin
9	3446	58.6	1163	21 AAB07361	Human CD11c protei

10	3446	58.6	1163	23 ABG61470	Human Beta2 integr
11	3444	58.6	1163	24 ABU07406	Protein differenti
12	3411	58.0	1161	16 AAR78166	Human beta-2 integ
13	3411	58.0	1161	18 AAW23049	Human beta-2 integ
14	3411	58.0	1161	19 AAW72825	Human alpha-d. Ho
15	3411	58.0	1161	19 AAW65089	Human Beta2-integr
16	3411	58.0	1161	19 AAW57491	Human Beta2-integr
17	3411	58.0	1161	20 AAW73342	Human alpha d clon
18	3411	58.0	1161	21 AAB07359	Human Beta2-integr
19	3411	58.0	1161	23 ABG61468	Human beta 2 integ
20	3395.5	57.8	1161	18 AAW23064	Human alpha-d deri
21	3395.5	57.8	1161	19 AAW72837	Human Beta-integr
22	3395.5	57.8	1161	19 AAW65106	Human Beta-integr
23	3395.5	57.8	1161	20 AAW73343	Human alpha d prote
24	3395.5	57.8	1161	21 AAB07376	Human alpha d prote
25	3395.5	57.8	1161	23 ABG61485	Human Beta2-integr
26	3234.5	55.0	1161	16 AAR78169	Rat alpha-d subuni
27	3232.5	55.0	1161	18 AAW23062	Rat beta 2 integr
28	3232.5	55.0	1161	19 AAW72824	Rat alpha-d #1. R
29	3232.5	55.0	1161	19 AAW60004	Rat alpha d polype
30	3232.5	55.0	1161	21 AAB07374	Rat alpha d protei
31	3232.5	55.0	1161	23 ABG61483	Rat Beta2-integrin
32	3225.5	54.9	1161	20 AAW65104	Rat beta-integrin
33	3225.5	54.9	1161	19 AAW73345	Rat alpha d integr
34	3222	54.8	1161	18 AAW23061	Mouse beta 2 integ
35	3222	54.8	1161	19 AAW72836	Mouse alpha-d #2.
36	3222	54.8	1161	19 AAW65103	Mouse beta-integr
37	3222	54.8	1161	19 AAW60003	Mouse alpha d poly
38	3222	54.8	1161	20 AAW73347	Mouse alpha d prote
39	3222	54.8	1161	21 AAB07373	Mouse Beta2 integr
40	3222	54.8	1161	23 ABG61482	Mouse alpha-d subu
41	3218	54.8	1161	16 AAR78168	Rat beta 2 integr
42	3211.5	54.7	1151	18 AAW23059	Rat alpha-d #2. R
43	3211.5	54.7	1151	19 AAW72834	Rat beta-integrin
44	3211.5	54.7	1151	19 AAW65101	Rat alpha d polype
45	3211.5	54.7	1151	19 AAW60001	

ALIGNMENTS

RESULT 1
AAW65090

ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

XX 28-SEP-1998 (first entry)

DE Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
lung inflammation; acute respiratory distress syndrome; CD11b subunit;
rheumatoid arthritis.

OS Homo sapiens.

XX US5728533-A.

PN 17-MAR-1998.

PD 07-JUN-1995; 95US-0485618.

XX 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 05-AUG-1994; 94US-0286889.

PR 21-DEC-1994; 94US-0362652.

XX (ICOS-) ICOS CORP.

PI Gallatin WM, Van DER VIEREN M;

DR WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-I diabetes

XX Example 5; Fig 1A-D; 106pp; English.

CC This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX Sequence 1153 AA;

SQ Query Match 99.8%; Score 5862; DB 19; Length 1153;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPOEIVAAQNRGLYQCDYSTGCEPI 60

DB 17 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPOEIVAAQNRGLYQCDYSTGCEPI 76

QY 61 RLQVPVEAVNMLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQOPK 120

DB 77 RLQVPVEAVNMLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQOPK 136

QY 121 FPEALRGCPQESDIAFLVDGSGIIPHDFRAKEPISTVMEOLKSKTFLSLMAYSEEF 180

DB 137 FPEALRGCPQESDIAFLVDGSGIIPHDFRAMEFVSTVMEOLKSKTFLSLMAYSEEF 196

QY 181 RIHFTKPEQNNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240

DB 197 RIHFTKPEQNNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 256

QY 241 TDGEKGDPLGYEDVTPADREGVIRYVGVGDAPFSEKSRQELNTVASKPRDHVQFN 300

DB 257 TDGEKGDPLGYEDVTPADREGVIRYVGVGDAPFSEKSRQELNTVASKPRDHVQFN 316

QY 301 NFPAKLTIONQLREKIFAETGTGSSSFHEHMSOEGFSAATSNGLLSTVGSYDWAQ 360

DB 317 NFPAKLTIONQLREKIFAETGTGSSSFHEHMSOEGFSAATSNGLLSTVGSYDWAQ 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSCPL 480

DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSCPL 496

QY 481 PRGORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDAIGAPEEDNRGAVLYF 540

DB 497 PRGORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDAIGAPEEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600

DB 557 HGTSGSGISPSHSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLVRVKAIMEFNPFEARNVFNCDQVVGKEAGEVRVCLHVQKSTRDRRLREGIQISVVT 660

DB 617 PVLVRVKAIMEFNPFEARNVFNCDQVVGKEAGEVRVCLHVQKSTRDRRLREGIQISVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 720

DB 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 736

QY 721 SLVGTPLSAFQNLRPVLAEDAQRULTALFPEKCKGNDNICODDLSTIFSPMSLDCLVWG 780

DB 737 SLVGTPLSAFQNLRPVLAEDAQRULTALFPEKCKGNDNICODDLSTIFSPMSLDCLVWG 796

QY 781 GPREFNVTVVRNDEGDSYRTQVTFFFPDLDSYRKVSTLQNRORSWRLACESASSTEV 840

DB 797 GPREFNVTVVRNDEGDSYRTQVTFFFPDLDSYRKVSTLQNRORSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900

DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

QY 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHQYVSNLGRSLPISLVFLVPV 960

DB 917 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHQYVSNLGRSLPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020

DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 FFGIQEENATLKGNSLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFFVRSQTET 1080

DB 1037 FFGIQEENATLKGNSLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFFVRSQTET 1096

QY 1081 KVEPPEVPNPLIVGSSVGGLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137

DB 1097 KVEPPEVPNPLIVGSSVGGLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX AC AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha_d integrin;

XX KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

XX KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

XX KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

XX KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

XX KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PF 16-NOV-1999; 99WO-US27139.

XX PR 16-NOV-1998; 98US-0193043.

XX PR 08-JUL-1999; 99US-0350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin MW, Van Der Vieren M;

XX DR WPI; 2000-387751/33.

XX PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

XX PT macrophage infiltration and reduce inflammation at central nervous

XX PT system injury sites -

XX PS Example 5; Fig 1; 270pp; English.

XX CC Integrins are a class of membrane-associated molecules that participate

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

XX
SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 21; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQSRVVVVGAPQEIIVANQKSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQSRVVVVGAPQEIIVANQKSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCTVHTQTCSENTYVVKGLCFPLFGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCTVHTQTCSENTYVVKGLCFPLFGSNLRQPOK 136

QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTFLSLMQYSEEF 196

QY 181 RIHFTTFEFQNNPRSLIKPITQLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240
DB 197 RIHFTTFEFQNNPRSLIKPITQLLGRTHATGIRKVVRELFINITGARKNAFKILVVI 256

QY 241 TDGSKFGDPLGYEDVIPADREGVIRYVIGVDGAFRSEKSKQELNTVASKPPRDRHVQIN 300
DB 257 TDGSKFGDPLGYEDVIPADREGVIRYVIGVDGAFRSEKSKQELNTVASKPPRDRHVQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTOTGSSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTOTGSSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQVSVCP 480
DB 437 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQVSVCP 496

QY 481 PRGQARWQCDVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLRVKAIMFNPREAVNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660
DB 617 PVLRVKAIMFNPREAVNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 676

QY 661 YDLALDSGRPHSRVNFENKSTRQTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVNFENKSTRQTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCLVVG 780

DB 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGDSYETQVTFPEPDLVSRYKSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGDSYETQVTFPEPDLVSRYKSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSESEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPFENSESEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVTXYLNFNTASNTSRVMQHQVQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSHGVTXYLNFNTASNTSRVMQHQVQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSQTE 1080
DB 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSQTE 1096
QY 1081 KVBPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGPPGAPQ 1137
DB 1097 KVBPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGPPGAPQ 1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX AC ABG61469;
XX XX 27-AUG-2002 (first entry)
XX DE Human Beta2 integrin alphaCD11b subunit.
XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX PN WO200230980-A2.
XX XX 18-APR-2002.
XX PD 15-OCT-2001; 2001WO-US32059.
XX PF 13-OCT-2000; 2000US-0688307.
XX PR (ICOS-) ICOS CORP.
XX FA Gallatin WM, Van Der Vieren M;
XX PI WPI; 2002-463260/49.
XX DR
XX XX
PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX Example 5; Page 191-194; 270pp; English.
XX The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha2 CDNA and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia. The present sequence is a Beta2 integrin alpha subunit sequence included for comparison with the Beta2 integrin alphaad protein sequences.

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 23; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAANQORSLVQCDSYSGCEPI 60
 DB 17 FNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAANQORSLVQCDSYSGCEPI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLSGNLRLQRPQK 120
 DB 77 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLSGNLRLQRPQK 136

QY 121 FPEALRGCPQEDSIAFLVDGSGSIIPHDFFRAKBFISTVMEQLKKSKTFLSLMOYSEEF 180
 DB 137 FPEALRGCPQEDSIAFLVDGSGSIIPHDFFRAKBFISTVMEQLKKSKTFLSLMOYSEEF 196

QY 181 RHFTFKPEQNNPRLSLKPTTOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
 DB 197 RHFTFKPEQNNPRLSLKPTTOLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256

QY 241 TGEKFGDPLGVEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPDRHVFQIN 300
 DB 257 TGEKFGDPLGVEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPDRHVFQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLLSTVGSYDWAG 360
 DB 317 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLLSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVLCAPRYOHIGLVAMPR 420
 DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVLCAPRYOHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCLP 480
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCLP 496

QY 481 PRQQRARWCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
 DB 497 PRQQRARWCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 616

QY 601 PVLVRVKAIMEFNPVARNVFECDQVWKGKAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
 DB 617 PVLVRVKAIMEFNPVARNVFECDQVWKGKAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDVPSPVLRNLF 720
 DB 677 YDLALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDVPSPVLRNLF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVWG 780
 DB 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVWG 796

QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 916

QY 901 QLELPVYAVYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLPISLVFLVPV 960
 DB 917 QLELPVYAVYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLPISLVFLVPV 976

QY 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1020
 DB 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1036

QY 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVTAEILFNDSVFTLLPGQGAFFVRSOTET 1080
 DB 1037 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVTAEILFNDSVFTLLPGQGAFFVRSOTET 1096

QY 1081 KVEPFEPVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGSGPPGAEPQ 1137
 DB 1097 KVEPFEPVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGSGPPGAEPQ 1153

RESULT 4
 AAU80252
 ID AAU80252 standard; Protein; 1153 AA.
 AC AAU80252;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human integrin 1 alpha-M subunit protein.
 XX
 KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
 KW inflammatory disease; autoimmune disorder; Crohn's disease;
 KW human immunodeficiency virus; HIV; myocardial infarction;
 KW Sjorgen's syndrome; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 499..500
 FT /note= "Encoded by GCG CAG AGG"
 FT
 XX WO200218583-A2.
 XX 07-MAR-2002.
 XX 31-AUG-2001; 2001WO-US27227.
 XX 01-SEP-2000; 2000US-229700P.
 XX (BLOO-) CENT BLOOD RES INC.
 XX Springer TA, Shimoaka M, Lu C;
 XX WPI; 2002-382964/41.
 DR N-PSDB; ABK50046.
 XX
 PT Modified integrin-I or integrin I-like domain polypeptide useful as an
 PT immunogen to produce antibodies specific to polypeptide, comprises a
 PT disulfide bond such that polypeptide is stabilized in a desired
 PT conformation
 XX
 PS Disclosure; Page 109-112; 112pp; English.
 XX
 CC This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilised in a desired conformation. The polypeptide of
CC the invention may have anti-inflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
CC infarction, Sjogren's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.
XX
SQ Sequence 1153 AA;

Query March 99.8%; Score 5862; DB 23; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDENAMTFOENARFGQSVVQLQGSRRVVGVAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTFOENARFGQSVVQLQGSRRVVGVAPOEIVAAQNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEFISTVMEQKKSKTFLPSLMQYSEEF 180
Db 137 PPEALRGCPQSDSIAFLVDGSGSIIIPHDPRMKEFVSTVMEQKKSKTFLPSLMQYSEEF 196
Qy 181 RIHFTFEFQNNPNSRLIPITOLLGRTHATGIRKVVRELFNITGARKNAFKILLI 240
Db 197 RIHFTFEFQNNPNSRLVLPITOLLGRTHATGIRKVVRELFNITGARKNAFKILLI 256
Qy 241 TDGKFKGDPGLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFKGDPGLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAAIIILNRVQSLVLAGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAAIIILNRVQSLVLAGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDAVLGYEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 497 PRGQARWQCDAVLGYEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDGLTMDGLVDLTVGAGHVILLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDGLTMDGLVDLTVGAGHVILLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

Qy 721 SLVGTPLSAGNLRPLVAEDAQRALFTALPFPFKNCGNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAGNLRPLVAEDAQRALFTALPFPFKNCGNDNICODDLSITFSFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 840
Db 797 GPREFNVTVVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 856
Qy 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 916
Qy 901 QLELPVKYAYVMVTVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAYVMVTVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNCSTAVCORIQCDIP 1036
Qy 1021 FFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1080
Db 1037 FFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1096
Qy 1081 KVEPFFVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKDMKSEGPPGAEPQ 1137
Db 1097 KVEPFFVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKDMKSEGPPGAEPQ 1153
RESULT 5
AA014428
ID AAO14428 standard; protein; 1153 AA.
XX AAO14428;
AC AAO14428;
XX
DT 03-MAY-2002 (first entry)
XX
DE Integrin Mac-1 alpha subunit.
XX
KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX
OS Unidentified.
XX
PN WO200204521-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US21805.
XX
PR 07-JUL-2000; 2000US-216600P.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PA (BLOO-) CENT BLOOD RES.
XX
PI Springer T;
XX
XX WPI; 2002-148167/19.
XX
PT New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders
XX
PS Example 1; Fig 1F; 90pp; English.
XX
CC The invention comprises structurally biased variant integrin inserted (I)
CC domain proteins, wherein the alterations to the protein occur in at least
CC

two noncontinuous regions. Specifically the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating: ischaemia/ reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin.

Query Match	99.8%	Score 5862;	DB 23;	Length 1153;
Best Local Similarity	99.3%	Pred. No. 0;		
Matches 1129; Conservative		7; Mismatches	1; Indels	0; Gaps 0;

Qy		1	FNLDTENAMTFOENARGFGSQSVOLQGSRVVVGAPOEIIVAAORGSLYOCDYSTGSCBPI	60
Dd		17	FNLDTENAMTFOENARGFGSQSVOLQGSRVVVGAPOEIIVAAORGSLYOCDYSTGSCBPI	76
Qy		61	RLOVPVEAVNWSLGLSLSLAATTSPOLLACGPVTHQTCSENTYYKGLCFPLSGNLRQQOK	120
Dd		77	RLOVPVEAVNWSLGLSLSLAATTSPOLLACGPVTHQTCSENTYYKGLCFPLSGNLRQQOK	136
Qy		121	PFEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTWMEQLKKSKTLFSLMQYSBEF	180
Dd		137	PFEALRGCPQEDSDIAFLVDGSGSIIPHDPRRMEKFVSTWMEQLKKSKTLFSLMQYSBEF	196
Qy		181	RHFHTFKGFQNPNPRSUKPITOLLGRTHATGIRKVVRLEFNITNGARKNAFKILILI	240
Dd		197	RHFHTFKGFQNPNRSVLVKPITOLLGRTHATGIRKVVRLEFNITNGARKNAFKILIVI	256
Qy		241	TDXEKGFDPLGYEDVIPADREGVIRYVIGVGDADRSEKSQBELNTVASKPRDRHVQIN	300
Dd		257	TDXEKGFDPLGYEDVIPADREGVIRYVIGVGDADRSEKSQBELNTIASKPRDRHVQVN	316
Qy		301	NPEALKUTONQUREKIFAIEGTQTGSSSSPFHEMSQEGPSAAITSNGPLLSTVGSYDWAQ	360
Dd		317	NPEALKTIQNQUREKIFAIEGTQTGSSSSPFHEMSQEGFSAAITSNGPLLSTVGSYDWAQ	376
Qy		361	GVPFLYTSKEKSTFINMTVRDSMDNDAYLGAAAAIILNRVQSILVLCAPRYQHIGLVAMER	420
Dd		377	GVPFLYTSKEKSTFINMTVRDSMDNDAYLGAAAAIILNRVQSILVLCAPRYQHIGLVAMER	436
Qy		421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCVPL	480
Dd		437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCVPL	496
Qy		481	PRGORARWCDDAVLYGEQGQPWRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAUYLF	540
Dd		497	PRGORARWCDDAVLYGEQGQPNRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAUYLF	556
Qy		541	HGTSGSGISPHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGQAQGHVLLRSQ	600
Dd		557	HGTSGSGISPHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGQAQGHVLLRSQ	616
Qy		601	PVLURVKAIMEFPNPREVARNVFECNQDVVKGEAGEVRVCLHVQKSTRDRLREQIOISVVT	660
Dd		617	PVLURVKAIMEFPNPREVARNVFECNQDVVKGEAGEVRVCLHVQKSTRDRLREQIOISVVT	676
Qy		661	YDALDLSGRPHSRVAFNETKNSTRTOTVLGITCETLKLQLPNCIEDBPVSPIVLRLNF	720
Dd		677	YDALDLSGRPHSRVAFNETKNSTRTOTVLGITCETLKLQLPNCIEDBPVSPIVLRLNF	736
Qy		721	SLVGTPLPSAFGNLRPVLJAEAOQLFTALFFPEKNCNGDNICODDLSTIFSFMSLDCLVWG	780
Dd		737	SLVGTPLPSAFGNLRPVLJAEAOQLFTALFFPEKNCNGDNICODDLSTIFSFMSLDCLVWG	796
Qy		781	GPREFNVTVTRNDGEDSYRTQVTFPPFLDIJSYRKVSTLNQRSQRSWRLACCASSSTEY	840

FT Modified-site 1045..1047
 FT /*label= putative N-glycosylation site
 FT Modified-site 1051..1053
 FT /*label= putative N-glycosylation site
 FT Modified-site 1076..1078
 FT /*label= putative N-glycosylation site
 FT region 1..16
 FT /*label= signal peptide
 FT region 1106..1134
 FT /*label= putative transmembrane region
 XX EP364690-A.
 PN
 XX
 XX 25-APR-1990.
 PD
 XX
 XX 17-AUG-1989; 89EP-0115159.
 PP
 XX
 XX 23-AUG-1988; 88US-0235353.
 PR
 PR 09-MAR-1989; 89US-0321239.
 XX
 XX (DAND) DANA FARBER CANCER INST INC.
 PA
 XX
 XX Springer TA, Corbi A;
 PI
 XX WPI; 1990-125938/17.
 XX N-PSDB; AAQ04043.
 DR
 XX
 XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 PT
 XX Disclosure; Page ?; ?pp: English.
 PS
 XX
 XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX
 SQ Sequence 1153 AA;
 Query Match 99.6%; Score 5852; DB 11; Length 1153;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTFQENARFGQSVQLOGSRVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARFGQSVQLOGSRVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVAVNMSLGLSAAATSPQLLACGPTVHQTCSNTYVYKGLCLFGLSGNLROQPOK 120
 DB 77 RLQVPVAVNMSLGLSAAATSPQLLACGPTVHQTCSNTYVYKGLCLFGLSGNLROQPOK 136
 QY 121 FPEALRCQPDSDIAFLDQSGSIIIPHDPRRAKEFTSTVNEQLKSKTILFSLMOYSEEF 180
 DB 137 FPEALRCQPDSDIAFLDQSGSIIIPHDPRRAKEFTSTVNEQLKSKTILFSLMOYSEEF 196
 QY 181 RIHFTFKFQNNPRSLKIPITQLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
 DB 197 RIHFTFKFQNNPRSLKIPITQLGRTHATGIRKVVRELFINITNGARKNAFKILILI 256
 QY 241 TDGEKFGDPLGYEDIVEADREGVIRYVIGVDGAFRSEKSKQELNVTASKPRDHVFQIN 300
 DB 257 TDGEKFGDPLGYEDIVEADREGVIRYVIGVDGAFRSEKSKQELNVTASKPRDHVFQIN 316
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQGFSAIITNSGPLLSTVGSYDWAG 360
 DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQGFSAIITNSGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 420
 DB 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 480
 DB |||||
 QY 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 496
 DB |||||
 QY 481 PRQORARWQCDVLYGEGQGPWGRFGAALTFLVDGVNGDKLTDVAIGAPGEDNRGAYVLF 540
 DB |||||
 QY 497 PRQORARWQCDVLYGEGQGPWGRFGAALTFLVDGVNGDKLTDVAIGAPGEDNRGAYVLF 556
 DB |||||
 QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 600
 DB |||||
 QY 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 616
 DB |||||
 QY 601 PVLVRKAIMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVOKSTRDRREGIOQSVVT 660
 DB |||||
 QY 617 PVLVRKAIMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVOKSTRDRREGIOQSVVT 676
 DB |||||
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDVPSIVLRNLF 720
 DB |||||
 QY 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDVPSIVLRNLF 736
 DB |||||
 QY 721 SLVGTPLSARGNLRPVLAEDAORLFTALPFPEKNCNDNTICODDLSITFSFMSLDCJLVG 780
 DB |||||
 QY 737 SLVGTPLSARGNLRPVLAEDAORLFTALPFPEKNCNDNTICODDLSITFSFMSLDCJLVG 796
 DB |||||
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRMLACESASSTEV 840
 DB |||||
 QY 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRMLACESASSTEV 856
 DB |||||
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 900
 DB |||||
 QY 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 916
 DB |||||
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQVNSLQORSIPISLVFLVPV 960
 DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQVQVNSLQORSIPISLVFLVPV 976
 QY 961 RLNQTVIWRDPQVTFSENLSSTCTKERLPSHSDFLAELRKAPVWNCISIAVCORIOCDIP 1020
 DB |||||
 QY 977 RLNQTVIWRDPQVTFSENLSSTCTKERLPSHSDFLAELRKAPVWNCISIAVCORIOCDIP 1036
 DB |||||
 QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGOGAFVRSQTET 1080
 DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGOGAFVRSQTET 1096
 QY 1081 KVEPFEPVNPPLPIVGVSSVCGLLLLALITAAALYKGLFFKRYQKDMMSGEGPPGAEPQ 1137
 DB 1097 KVEPFEPVNPPLPIVGVSSVCGLLLLALITAAALYKGLFFKRYQKDMMSGEGPPGAEPQ 1153
 RESULT 7
 AAR071120
 ID AAR071120 standard; protein; 1163 AA.
 XX
 AC AAR071120;
 XX
 DT 25-MAR-2003 (updated)
 DT 05-FEB-1991 (first entry)
 XX
 XX p150.95 alpha subunit encoded by clone lambda47.
 XX p150.95 leucocyte adhesion receptor alpha-subunit;
 XX hairy cell leukaemia; rhinovirus.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 61..63
 FT /label-glycosylation site
 FT Modified-site 89..91
 FT /label-glycosylation site
 FT Modified-site 385..387
 FT /label-glycosylation site

Human Beta-integrin CD11c subunit protein.

Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
lung inflammation; acute respiratory distress syndrome; CD11c subunit;
rheumatoid arthritis.

Homo sapiens.

US5728533-A.

17-MAR-1998.

07-JUN-1995; 95US-0485618.

07-JUN-1995; 95US-0485618.

23-DEC-1993; 93US-0173497.

05-AUG-1994; 94US-0286889.

21-DEC-1994; 94US-0362652.

(ICOS-) ICOS CORP.

Gallatin WM, Van DER VIEREN M;

WPI; 1998-206565/18.

Screening assay for modulators of integrin binding - using
immobilised or labelled alpha-d polypeptide, useful for, e.g.
treating type-I diabetes

Example 5; Fig 1A-D; 106pp; English.

This sequence represents a human beta-integrin CD11c subunit which is
used to describe a method for identifying compounds that modulate the
interaction of the beta-integrin alpha-d subunit with a binding partner
of alpha-d which involves contacting an alpha-d polypeptide with an
alpha-d binding partner, one of which is immobilised and the other of
which is labelled, in the presence of a test compound, and determining if
the compound affects binding between the alpha-d polypeptide and alpha-d
binding partner, where the alpha-d polypeptide is alpha-d or its fragment
comprising the cytoplasmic, transmembrane or extracellular domain of
alpha-d. Compounds that modulate alpha-d binding could be used to treat
diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
asthma, psoriasis, lung inflammation, acute respiratory distress
syndrome and rheumatoid arthritis.

Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 19; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1e-278;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

1 FNLTENAMTQENARGQSVVQLGSRVVGAPQEIIVANQKSLVQCDYSTGSCPEI 60
20 FNLTDELTAFRVDGAGDSVQVANSVWVVGAPQKIIAANQIGLYQCYSSTACEPI 79
61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHCTCSENTYVKGCLFLGSLNRQPOK 120
80 GLQVPPEAVNMSLGLSLASTTSPQLACGPTVHCEGRNMYLGLCLPLGLPT--QLTOR 137
121 FPEALRGCPQBDSDIAFLVDSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMQYSEEF 180
138 LPVSRQECPRQEQDIFLIDSGSISSENFMNFRVAVISQFQSTQSLMQPSNKF 197
181 RIHTFFKEFQNNPNSRIKIPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
198 QHTFFEFRRSTNPLSLASVHQLQGYTATATQNVVHRLFHASYGARRDAIKILIVI 257
241 TDGEKFGDPLYEDVIEADREGVIRYVIGVDFAFRSEKRSQELNTVASKPRPHVQIN 300
258 TDGKKGDSLDYKDVIPWADAAGTIRYVIGVLAFRNNSWKLNDIASKPSQEHIFKVE 317
301 NFEALKTIONQLREKIFAIEGTQTCGSSSSFEHMSQEGFSAAITNSGPLLSTLVGSYDWAG 360

Db DPALDKDQIQOLKEKIFAIEGTETETISSSSFELEMAQEGSAVFTDPDGVLGAVGFSWGS 377
Qy GVFLYTSKESKSTFINTRVDSMDNDAYLVGAAAIILRNRYQSLVGLGAPRYOHIGLVAMER 420
Db GATLYPPNMGPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVGLGAPRYOHIGKAVIFI 437
Qy QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIIGAPHYYEOTRGGOVSCPL 480
Db QVSRQWRMKAIEVIGTOIGSYFGASLCSVDVDTGSTDVLIIGAPHYYEOTRGGOVSCPL 497
Qy PRGORARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYLFP 540
Db PRGWR-RWWDVLYGEGHPWGRFGAALTVLGDVNGDKLTDVVGIGAPGEEENRGAYLFP 556
Qy HGTSGSGISPSHSORIASKLSPLRYFGOSLGGQDLTMDGLVDLTVGAGQHVLLLRSQ 600
Db HGVLGPSISPSHSORIASKLSRLQYFGALSGQDLTQDGLVDLAVGARGQVLLLRTR 616
Qy PVLKVKAIMBPNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDRREGIQSVVT 660
Db PVLWVGVSQMFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLGSRLQSSVT 676
Qy YDLALDSGRPHSRVAVNETKXSTRQTQVLTGLTQTCETLKLQLPNCIEDVPSPVLRLNF 720
Db LDALAPGRLSRAIFQETKNSLSRVVLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736
Qy SLVGTPLSARGLRPLVAEDAQRLEFALPFPEKNCNDNICODDLSITFSPMSLDCLAVG 780
Db TLVGKPLLAFLRPLMLAALAAQRYFTASLPFEKNCAGADHICQDNLGISFSFPGKLSLVG 796
Qy GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSQSRWRLACESASTEV 840
Db SNLENAEVMVNDGEDSYCTTTFSHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
Qy SGALKSTSCSINPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEP 900
Db SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
Qy QLELPVKYAVVMVTHSGVSTKYNLFTAS-ENTSRVMOHOYVSNLQORSIPISLVFLVP 959
Db QLELPVKYAVYVSSHEQTKYLNFSSEKESHVAMHRYQVNNLQORLPVNSINFWVP 974
Qy VRLNQTIVMDRQVPTSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORICDI 1019
Db VELNQEAVMVDVSHVHPQNSLRCSEKIAPPASDFLAHQKXPVLDSCAGCLRPRCDV 1034
Qy PFGIOBEFNATLKGNSLFDWYIKTSHNHLILVSTABILFNDVSFTLLPGGAFVRSQTE 1079
Db PSFSVQBELDFTLKGNSLFGWVRQILQKKVSVSVABIIIFDTSVYSQLPQGAEPMAQTI 1094
Qy TKVEPPEVPNPLPLIVGSSVGGILLLALITAAALYKLGFFKRYOKDMWSE 1128
Db TVLEKYYKHNPIPLIVGSSIGGULLLALITAVLYKVGFFKRYOKEMMEE 1143
RESULT 9
AAB07361
ID AAB07361 standard; Protein; 1163 AA.
XX AAB07361;
AC AAB07361;
DT 17-JAN-2001 (first entry)
DE Human CD11c protein sequence.
KW Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX

OS Homo sapiens.
 PN WC0200029446-A1.
 XX
 PD 25-MAY-2000.
 XX
 PF 16-NOV-1999; 99WO-US27139.
 XX
 PR 16-NOV-1998; 98US-0193043.
 PR 08-JUL-1999; 99US-0350259.
 XX
 PA (ICOS-) ICOS CORP.
 PI Gallatin MW, Van Der Vieren M;
 XX
 DR WPI; 2000-387751/33.
 XX
 PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous
 PT system injury sites
 XX
 PS Example 5; Fig 1; 270pp; English.
 XX
 CC Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This
 CC sequence was used in an alignment to identify a novel beta2 integrin
 CC alpha subunit: alpha_d (AA060014 and AAB07359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha_d. The
 CC Alpha_d gene and protein may be useful in therapy for diseases linked
 CC to aberrant alpha_d function e.g. Type I diabetes, atherosclerosis,
 CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
 CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
 CC adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal
 CC antibodies may be used in the inhibition of macrophage infiltration at
 CC the site of a central nervous system injury. The monoclonal antibodies
 CC can also be used to detect and diagnose Crohn's disease.
 XX
 SQ Sequence 1163 AA;
 Query Match 58.6%; Score 3446; DB 21; Length 1163;
 Best Local Similarity 60.8%; Pred. No. 1e-278;
 Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFQENARGFGOSVVOLOQSRVWVGAPQEIIVANORGSLVQCDYSTGSCPEI 60
 DB 20 FNLDTBELTAFRVDSDAGSDSVVOYANSWVVGAPQKIIAAQIIGGLYQCYSTGACEPI 79
 QY 61 RLQVPVAVNMSLGLSLAAATTSPPQLLACGPTVHQTCSNTYVYKGLCFPGSNLRQOPQK 120
 DB 80 GLQVPPAVNMSLGLSLASTTSPQLLACGPTVHHECGRNWYLTGLCFLLGPT-QLTQR 137
 QY 121 FPEALRCGPQEDSDIAELVDCSGSIIPHDFFRAKEPFTSTVMEQLKKSKTLPFLMQVSEEF 180
 DB 138 LPVSRQECPRQEQDIFELIDSGSGSISSRNFEATMNFVRAVISQFQRPSTQFSLMQFSNKF 197
 QY 181 RIHETTFEQQNPNSRLIPITQLLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
 DB 198 QTHETTFEQRRTSNPLSLASVHQLOQFTYTATAIQNVVHRLFASVAGARRDAIKILIVI 257
 QY 241 TDGKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELNVTASKPPRDHVPQIN 300
 DB 258 TDGKKGESLDYKDVIFWADAAGIIRVIAUGLAFQFNNSWKNELDIASKPSQEHIFKVE 317
 QY 301 NFEALKTIONQLREKIFAEIGTQTGSSSSFEHENSQEGFSAATTSNGLPLSTVGSYDWAG 360
 DB 318 DFDALKDIONQLREKIFAEIGTQTGSSSSFEHENSQEGFSAVFPDGPVLGAVGSFTWSG 377
 QY 361 GVFLYTSKESKSTFNMTTRVDSMDNDAYLGVAAAILLNRVQSLVGLGAPRQIHGLVAMFR 420

DB 378 GAFLYPPNNSPTFINMSQENMDRSDSYLGSYSTELALWKGVSQSLVGLGAPRQIHGKAVIFI 437
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYEQTGGQVSVCP 480
 DB 438 QVSRQWRMKAIEVIGTIGSYFGASLCSVDVDTGSDTLVILGAPHYEQTGGQVSVCP 497
 QY 481 PRGQARWQCCDAVLYGEOGQPWGRFGAALTVLGDNVNGDKLTDVAIGAPEEDNRGAVLYF 540
 DB 498 PRGWR-RWVCDVLYGEOGHPWGRFGAALTVLGDNVNGDKLTDVVGAPGEBENRGAVLYF 556
 QY 541 HGTSGSIGSPHSQRIAGSKLSPLRQYFGQSLSGGODLTMDGLVDLTVGAAQHVLRLRSQ 600
 DB 557 HCVLGFSPHSQRIAGSKLSPLRQYFGQSLSGGODLTMDGLVDLAVGARGQVLLLRTR 616
 QY 601 PVLRYKATMEFNPREVARNVPECNDVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
 DB 617 PVLVGVSNQFIPAEIPRSAFECEQVSVSEQTLVQSNICLYIDKSKNLLGSRDQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
 DB 677 LDLALAPGLSPRAIFQETKNRSLSRVRVLGLKHCENFNLLPSCVEDSVPIILRLNF 736
 QY 721 SLVGTPLSAFGNLRPLVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSMSLDCLVVG 780
 DB 737 TLVKGPLLAFLNLRPMLAALAQRFTASLPPEKNCADHICQDNLGIGSPFPLKSLVVG 796
 QY 781 GPREFNVTVVNDGEDSDYRTQVTFPFLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
 DB 797 SNLELNAEVMVWMDGEDSYGTTITFHPAGLSRYVVAEGQKQGLRSLHLTC--CSAPVG 854
 QY 841 SCALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB 855 SGTWSTSCRIINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFATAS-ENTSRVMQHYQVSNLQKSLPISLVFLVP 959
 DB 915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEKESHVAMHYQVNNLQORDLPVSNFWVP 974
 QY 960 VRLNQTVIWRQVTFSENLSTCHTKERLPSHDFLAELRKAPVNVNCSINVCQICQDI 1019
 DB 975 VELNQEAVMVDEVSHPQNPRLCSSEKIAPPASDFLAHQINQVLDSCSIAGCLFRCDV 1034
 QY 1020 PFFGLOEENATLKGNLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1079
 DB 1035 PSFSVOEELDTLKGNLSEGWVRQILQKKVSVVSAEIIFDTSVTSQLPQCAFRAQTI 1094
 QY 1080 TKVEPPEVPNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKQMMSE 1128
 DB 1095 TVLEKYKVHNPIPLIVGSSIGLLLALITAVLYKVGFFKQYKEMWEE 1143
 RESULT 10
 ABG61470
 ID ABG61470 standard; Protein; 1163 AA.
 XX
 AC ABG61470;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human Beta2 integrin alphaCD11c subunit.
 XX
 KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
 KW LAD; leukocyte adhesion deficiency; inflammation response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX

PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Jay G;
XX WPI; 2003-058520/05.
DR
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
PS Claim 1; Page 225-228; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;

Query Match 58.6%; Score 3444; DB 24; Length 1163;
Best Local Similarity 60.7%; Pred. NO. 1.5e-278;
Matches 685; Conservative 139; Mismatches 299; Indels 6; Gaps 4;

QY 1 FNLDTEANMTFQBNARGFGQSVVQLOGSRVVVGAPOEIVAAANGRSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTRAFRVDGAGFSDSVVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHQTCSNTYVYKGLCFGLGSLNLRQPOK 120
DB 80 GLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHQTCSNTYVYKGLCFGLGSLNLRQPOK 137
QY 121 FPAALGCCPQEDSDIAFLVDGSGSIIPHPRRAKEFISTVMEQLKSKTLFSLMZYSEEF 180
DB 138 LPVSRQCEPRQEQDIFVLIDGSGSISSRNFRATVWVFRVAVISQFQPSQFSLMQFSNKF 197
QY 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240

DB 198 QTHLTFTFEFRRTSNPLSLASVHQLQOFTYTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDCEKFGDPLGYEDVIPPEDAREGVIYVIGVGDFAFRSEKSRQBELNTVASKPRDHVQIN 300
DB 258 TDCKEGEDTLDYKDVIFPMADAAGIIRYAGVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
QY 301 NFPAKTIQNLREKIPAEIGTOTGSSSSPEHEMSQEGFSAALTSNGPLLSVTGSDVWAG 360
DB 318 DFDALDKIQTLREKIPFPIEGTETSSSPELEMAQEGFSAVTPDGPVLGAVGSEFTMSG 377
QY 361 GVELYTSKEKSTFINNTRVDSMDNDAVLGYAAAILRNRVQSLVGLAPRYQHLGLVAMFR 420
DB 378 GAFLYPPNMGPTFINNSQENVDNRDSYLGSTELALWGVQSLVGLAPRYQHTGKAVIFT 437
QY 421 QNTGMWESNANVKGITQIGFYGASLCSVDVDSNGSTDVLVIGAPHYYEQTQGGQSVCPPL 480
DB 438 QVSRQWRMKAETGTQIGSYFGPSLCSVDVDSGSTDVLVIGAPHYYEQTQGAQSVCPPL 497
QY 481 PRGORARQCDVLYGQGOQPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLFL 540
DB 498 PRGWR-RWMCDAVLYGQGHQWGRFGAALTVDGVNGDKLTDVVIGAPGEENRGAYLFL 556
QY 541 HGTSGSGISPSHSORIASGLSPLOYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGVLGSPISPSHSORIASGLSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLLRTR 616
QY 601 PVLKVAIMEFNPREVARNVPECNDQVVGKEAGEVRVCLHVQKSTRDRREGIQSVVT 660
DB 617 PVLWGVGSMQFIPAEIPRSAFECEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDGRPHSRVAVNETKNSRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 LDALDGRSLSPRATFOETKNSRVRVLGLKACENFNLLPSCVEDSVTPITRLNF 736
QY 721 SLVGTPLSAFQNLBPVLAEDAQRFTALFPEKCKGNDNICODDLSITFSPMSLDCLVVG 780
DB 737 TLVGKPLLAFLNRLPMLAADQRYFTASLPEKCKGADHICQDNLGIFSFPGLKSLVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWRLACESASSTEV 840
DB 797 SNLEUNAEVWVNDGEDSYGTITFSPHAGLSYRYVAEGQKQGLRSLHLTCDAPVG- 854
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRINHLIFRGAQITFLATPDVSPKAVLGDRLLLTANVSENNTPRTSKTF 914
QY 901 QLELPVYAVYVWVTSCHGVSTKYLNTAS-ENTSRVMOHOYOVSNLQORSPLISVFLVP 959
DB 915 QLELPVYAVYVWVSSHEQFTKYLNFSESEKESHVAMHYOVNMLGQDLPVSNFWVP 974
QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDI 1019
DB 975 VELNQEAVMQVDEVSPLQNPSPSRCSSEKIAGPSDFLAHIQKNPVLDCSIACCLFRCDV 1034
QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSTPLPGQAFVRSQTE 1079
DB 1035 PSFSVQEEELDFTLKGNLSFGVVRQILQKKVSVVAEITFTDTSVYSQLPQGEAFRAQT 1094
QY 1080 TKVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKQDMSE 1128
DB 1095 TVLEKYKVHNPETPLIVGSSIGLLLLALITAVLYKVGFQKQYKEMMEE 1143

RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.

KW	Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.	
KW		
XX		
OS	Homo sapiens.	
XX		
XX	Location/Qualifiers	
FT	17...1108	
FT	/note= "extracellular domain"	
FT	150...352	
FT	/note= "this region is homologous to the insert common to CH1a,b,c and may be a site for interaction with ICAM family proteins"	
FT		
FT	Binding-site	
FT	465...474	
FT	/note= "putative cation binding site"	
FT	518...527	
FT	/note= "putative cation binding site"	
FT	592...600	
FT	/note= "putative cation binding site"	
FT	1109...1128	
FT	/note= "transmembrane region"	
FT	1129...1161	
FT	/note= "cytoplasmic domain"	
XX		
XX	W09517412-A1.	
PN		
XX	29-JUN-1995.	
XX		
XX	21-DEC-1994; 94WO-US14832.	
PD		
XX		
XX	05-AUG-1994; 94US-0286889.	
PR		
XX	23-DEC-1993; 93US-0173497.	
XX		
XX	(ICOS-) ICOS CORP.	
PA		
XX		
XX	Gallatin WM, Van Der Vieren M;	
PI		
XX	WPI; 1995-240603/31.	
DR	N-PSDB; AAQ91712.	
XX		
XX	Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential antiinflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.	
PT		
PT	Claim 7; Page 82-87; 172pp; English.	
XX		
XX	A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha-TM1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.	
CC	The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells.	
CC		
XX		
SQ	Sequence 1161 AA;	
	Query Match 58.0%; Score 3411; DB 16; Length 1161;	
	Best Local Similarity 59.4%; Pred. No. 8.9e-276;	
	Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;	
Qy	1 FNLDTENAMTFOENARGGOSVQLOGSRVVVGAPQETIVANORGSLYQCDYSTGSCPEI 60	
Db	17 FNLDEVEPTIQEDAGFGQSVQFGSRLVVGAPLEVVAANQTRGLYDCAATGMCPQI 76	
Qy	61 RLOVPVEAVNMSLGLSAAATTPPQLLACGPTVHTQTCSENTYVKGCLCFLFGSNLRQOPQK 120	
Db	77 PLHIRPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSKSGCLLIGSRW-ELIQT 135	
Qy	121 PPEALRCQDSDIAFLVDGSGSIIPHDPRAKEFISTVNEQLKKSCTLPSLMQYSEEF 180	
Db	136 VPDATPCPHQEMDIVFLIDGSGIDQDNFNQMGKGFVQAVMQFEGDTLFLALMQYSNLL 195	
Qy	181 RIHETFEFQNNPNSRIKPIITOLLGTHTATGIRKVVRELFINITGARKNAFKILLI 240	
Db	196 KIHFTFTQFTSPSQQLVDPDIVQLKGLTFTATGILTVTLQFLHHKNGARKSAKKILIVI 255	

RESULT 13
AAW23049

ID AAW23049 standard; Protein; 1161 AA.

XX AAW23049;

AC AAW23049;

XX 24-FEB-1998 (first entry)

XX Human beta 2 integrin alpha d subunit.

XX Beta 2 integrin alpha d subunit; human; cell migration;
XX cell adhesion; phagocytosis; diabetes; arteriosclerosis;

Qy	241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVDAFRSEKSEKQELNLTNTVASKPPRDHVFQIN 300	
Db	256 TDGOKYKPLEYSDVIPQAEKAGIIRYAIVGHAFQGTARQELNTTSSAPPDQHVFKVD 315	
Qy	301 NFPAKTIQNLREKIPAIETGTQSSSSSEHEMSQEGFSAATSNQPLISTVGSYDWAG 360	
Db	316 NFAALGSIQLOKEIKIYAVEGTQSRASSSPQHEMSQEGFSTALTMDGLFLGAVGFSWSG 375	
Qy	361 GVFLYTSKEKSTFINMTRVDSMDNDAVLGVAAAIILNRRVQSLVIGAPRYOHIGLVAMFR 420	
Db	376 GAFLYPPNMSPTFINNSQENVMRDSYLGSTELALWKGVQNLVLGAPRYOHTGKAVIFT 435	
Qy	421 QNTGMESNANVRKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQSVSCPL 480	
Db	436 QVSRQWRKKAETGTQIGSYFGASLCSVDVDSNGSTDLLIIGAPHYYEQTGGQSVSCPL 495	
Qy	481 PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLFL 540	
Db	496 PRGORVQWQCDAVLRGEQGHPRFGAALTVLGDVNEDEKLIDVAIGAPGQENRGAYLFL 555	
Qy	541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQGVLLLRSQ 600	
Db	556 HGASEGISPSHSQRIAGSKLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLLRSL 615	
Qy	601 PVLRLVKAIMEFNPREVARNVFECDQVYKGEAGEVRVCLHVOKSTRDLRREGIOQSVVT 660	
Db	616 PVLKVGVMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673	
Qy	661 YDLALDSGRPHSRAVFNETKNTRRQTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720	
Db	674 FDLALDPGLTSAIFNETKNPTLTRRKTGLGLHCETLKLKLLPDCVEDVVSPIILHLNF 733	
Qy	721 SLVGTPLSAFGLNLRPVLAEDAORLFTALPFPEKCNKNDNICODDLSITFPMSLDCLVVG 780	
Db	734 SLVREPIPSQNLRPVLAAGSQDLFTASLPFKNCGQDGLCEGDLVTLSFGSLQTLTVG 793	
Qy	781 GPREFNVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 840	
Db	794 SSLELNVITVMNAGEDSYGTVSVLYYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852	
Qy	841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900	
Db	853 EG-LRSSRCNVNHPHFHEGSGNGTFIVTFDVSYKATLGDRLMLRASASSENKASSKATF 911	
Qy	901 QLELPVKYAVYVMVTSHGVSTKYLN--TASENTSRYVMQHOYOVSNLCORSIPISLVFLVP 959	
Db	912 QLELPVKYAVYTMISROESTKIFNATSEKMKKEAHRVYRNLSQORDLAISINFWVP 971	
Qy	960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDI 1019	
Db	972 VLLNGVAVMDVMEAPQSGL--PCVSEKPKPQHSDFLTQISRSPMLDCCSIADCLQPRCDV 1029	
Qy	1020 PFTGIEQEFNATLKNLSFDWYIKTSHNHLLIVSTBAILFNDSVFTLLPGGAFVRSQTE 1079	
Db	1030 PSFSVQBELFTLKNLSFGWVRETLQKKVLVSVABEITDTSVYSLQPLQGEAFMQAME 1089	
Qy	1080 TKVEPEVPNPLPIVGVSSVGGLLALLALITAAALYKLGFFKRYQKQWMMSE 1128	
Db	1090 MVLEDEEVYNAIPIIMGSSVGALLLALITATLYKLGFFKRYKEMUELD 1138	

KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.
 XX Homo sapiens.
 OS US831029-A.
 PN 03-NOV-1998.
 PD 07-JUN-1995; 95US-0482293.
 XX 07-JUN-1995; 95US-0482293.
 XX 23-DEC-1993; 93US-0173497.
 PR 05-AUG-1994; 94US-0286889.
 PR 21-DEC-1994; 94US-0362652.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1998-609318/51.
 DR N-PSDB; AAV67281.
 XX Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 PT immunohistochemical analysis
 XX Example 5; Column 61-66; 106pp; English.
 XX The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2). Antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX SQ Sequence 1161 AA;
 Query Match 58.0%; Score 3411; DB 19; Length 1161;
 Best Local Similarity 59.4%; Pred. No. 8.9e-276;
 Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
 QY 1 ENLDTENAMTFQENARGFGOSVVOLOGSRVVVGAPQEIIVAAQORGSLYOCDSYSGSCEPI 60
 DB 17 FNLDDVEPTIFQSDAGFGOSVVOFGSRVLVVGAPLEVAANOTGRLYDCAATGMCPOI 76
 QY 61 RLQVPVEAVNMSLGLSAAITSPPLQACGPTVHQTCSNTYVVKGLCFGLFGSNLRQOPQK 120
 DB 77 PLHIREAVNMSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIIQT 135
 QY 121 FPEALRGCPQEDSDIAPLVGSGSIIPHDFRRAKEFISTVNEQLKXSKTLFSLMQYSSEF 180
 DB 136 VPDATPECPHQEMDIVFLIDGSGSIDQNDPNQMGFVQAVMGQFEGDTLFLALMQYSNLL 195
 QY 181 RIHFTPEFQNNPRSLIKPIITQLLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
 DB 196 KIHFTTFQRTSSQOSLVDPIVOLKGLTFTATGILTVTQTLFHHKNGARKSAKKILIVI 255
 QY 241 TDCEKFDGLGYEDVPEADREGVIRVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
 DB 256 TDQOKYKDPLEYSDVIPAQKAGIIRYALGVGHAFQPTARQELNIISSAPPDHFVKVD 315
 QY 301 NFPAKTIQNLREKIFAIEGTQTGSSSPSEHEMSQEGFSAAITSGNPLLSVTGSDWAG 360
 DB 316 NFAALGSIQOLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMQDLFLGAVGSPFSWSG 375

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
 DB 376 GAFLYPPNMSPTFINNSQENVMDRDSYLGSTELALWKGVQNLVLGAPRYQHTGKAVFT 435
 QY 421 QNTGMWESNANVKGTOIGAYFCASLCSVDVDSNGSTDLVIGAPHYETRGQVSVCP 480
 DB 436 QVSRQWRKKAQVETGTQIGSYFGASLCSVDVDSNGSTDLVIGAPHYETRGQVSVCP 495
 QY 481 PRQORARQCDVAVLYGEOQPMGRFGAALTVLGVDVNGDKLTDVAIGAPGEENRGAVYLF 540
 DB 496 PRQORVQWQCDVAVLRGEQGHPMGRFGAALTVLGVDVNGDKLTDVAIGAPGEENRGAVYLF 555
 QY 541 HGTSGSISPSHSQRIAGSKLSPRIQYFQSGSUGGODLTMDGLVDLTVGAGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRIQYFQALSGGQDLTQDGLMDLAVGARGVLLRSL 615
 QY 601 PVLRYKATMEFNPVARNVFCNDQVVKGEAGEVVRCLHVOKSTRDLREGQIQSVVT 660
 DB 616 PVLKVGAMRFSPVEVAKAVYRCWBEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVFNENETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
 DB 674 FDLALDPGLTSRAIFNETKNPTLTKRTGLGIHCETLKLPLDCVEDVSPVILHLNF 733
 QY 721 SLVGTPLSAFNLPRVLAEDAQRLLFTALFPFPEKNCNDNICQDDLSITFSPMSLDCLVVG 780
 DB 734 SLVREPIPSQNLPRVLA VGSQDLFTASLPFPEKNCQDGLCEGLGVTLFSGLQTLTVG 793
 QY 781 GPREFNVTYVNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSWLACESASSTEV 840
 DB 794 SSLELNVIYVWVNGEDSYGTVSYLYPAGLSHRRVSGAKQPHOSALACETV--PTED 852
 QY 841 SGALKSTSCSINHPIFFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 DB 853 EG-LSSRCSVNHPIFHEGSGNCTFIVTDPVSKATLGDRLMLRASSENKASSSKATF 911
 QY 901 QLELPKYAVYVMVTVSHGVSTKYLNF-TASENSTRVMQHOYQVSNLQSRSLPISLVFLVP 959
 DB 912 QLELPKYAVYTMISRQEESTKYFNATSDKKMKEAEHRYRVNLSQDLAISINFWVP 971
 QY 960 VRLNQTVIWDPROVTFSENLSSTCHTKEPLPSHDFLAELRKAPVNVCSIAVCQICDI 1019
 DB 972 VLLNGVAVWVWMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQPCDV 1029
 QY 1020 PFEGIQEEFNATLKGNLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1079
 DB 1030 PFSVQEEELDTLKGNLSEGWVRETLQKVLVSVVAEITFTSVYSQZFGQAFMRAQME 1089
 QY 1080 TKVEPPEVNPPLIVGSSVGGLLIALITALYALYKLGFFKRYQKDMWSE 1128
 DB 1090 MVLEDEEVYNAIPPIINGSSVGALLLALITATLYKLGFFKRYKEMLED 1138
 RESULT 15
 AAW65089
 ID AAW65089 standard; Protein; 1161 AA.
 XX AAW65089;
 AC AAW65089;
 XX 28-SEP-1998 (first entry)
 DT Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 OS Homo sapiens.
 XX Key
 FH Peptide
 FT 1..16

```

FT Protein /label= signal
FT 17..1161
FT /note= "Beta-integrin alpha-D subunit"
FT Domain 17..1108
FT /label= extracellular
FT Domain 150..352
FT /note= "homologous to insertion domain of CD11 subunits"
FT Domain 1129..1161
FT /label= cytoplasmic
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-0485618.
XX
PR 07-JUN-1995; 95US-0485618.
PR 23-DEC-1993; 93US-0173497.
PR 05-AUG-1994; 94US-0286889.
PR 21-DEC-1994; 94US-0362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin WM, Van DER VIEREN M;
XX
DR WPI; 1998-206565/18.
DR N-PSDB; AAV3S236.
XX
PT Screening assay for modulators of integrin binding - using
PT immobilised or labelled alpha-d polypeptide, useful for, e.g.
PT treating type-I diabetes
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a novel human beta-integrin alpha-d subunit.
CC This sequence is used in a method for identifying compounds that modulate
CC the interaction of alpha-d with a binding partner of alpha-d which
CC involves contacting an alpha-d polypeptide with an alpha-d binding
CC partner, one of which is immobilised and the other of which is labelled,
CC in the presence of a test compound, and determining if the compound
CC affects binding between the alpha-d polypeptide and alpha-d binding
CC partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress
CC syndrome and rheumatoid arthritis.
XX
SQ Sequence 1161 AA;
Query Match 58.0%; Score 3411; DB 19; Length 1161;
Best Local Similarity 59.4%; Pred. No. 8.9e-276;
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
QY 1 FNLDTENAMTFQENARGFQGVVQGGSRVVGAPQEIIVAAHQSLYQCDYSTGSCPEI 60
DB 17 FNLDVEEPTIFEDAGGFGQSVVQGGSRVVGAPLEVVAAHQTLRVDCAATGWCQPI 76
QY 61 RLQVPVEAVNMSGLSLAATSPQQLACGPTVHQTCSNTYVKGCLFLGSLNRQQPQK 120
DB 77 PLHIRPEAVNMSGLTLAASNGSRLLACGPTLHRVCGENSYSKGSCLLGRW-ELIQT 135
QY 121 FEARLGGCPQESDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKSKTFLSLMOYSEF 180
DB 136 VPDATPECPHQEMDIVFLTDGSGSIDQDNFNQMGFVQAVMGQFEGTDTLFLALMOYSNLL 195
QY 181 RIHETFKFQNNPNRSLIKPTQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
DB 196 KIHFTTQRTSPSQSLVDPIVLQKGLTFTATGILTIVTQLFHHNGARKSAKILIVI 255
QY 241 TDGEKFGDPLGVEDYIPEADREGVIRYIVGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 256 TDGQYKDPLEYSVDVIPAQEKAGIIRYAIYGVGHATQGPARTQELNTISSAPPQDHVKVD 315

```

```

QY 301 NPEALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFAAITSNGPLLLSTVGSYDWAG 360
DB 316 NPAALGSIQLOLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFGAVGCSFWSG 375
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 376 GAFLYPPNMGPTFINMSQENVMDRDSYLGYSSTELALWKGVQNLVLGAPRYOHTGKAVFT 435
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGGOVSCPL 480
DB 436 QVSRQWRKKAEEVGTQIGSYFGASLCSVDVDSGSDTLIILIGAPHYYEOTRGGOVSCPL 495
QY 481 PRQORARWQCDAVLYGEOGQOPWGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAYVLF 540
DB 496 PRQORVQWQCDAVLYGEOGQHPWGRFGAALTVLGVDVNEKLDLIDVAIGAPEGQNRGAYVLF 555
QY 541 HGTSGSISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAGHVVLLRSQ 600
DB 556 HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLRSL 615
QY 601 PVLRYKAIMFENPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTDRLRREGIOISVVT 660
DB 616 PVLKVGVMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSVR 673
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQPNCIEDPVSPVILRLNF 720
DB 674 FDLALDPGRITSRRAIFNETKNPTLTRKTLGLGHCEHLKLLPDCVEDVSPILHLNF 733
QY 721 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPPEKNGNDNICQDLSITFSFMSLDCLVWG 780
DB 734 SLVREPISPQNLRPVLAVGSQDLFTASLPPEKNGQDGLCEGDLGVTLSFGSLQTLTVG 793
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSTYRKVSTLQNRQSRWRLACESASSTEV 840
DB 794 SSLELVIVITVMNAGEDSYGTWSLYYPAGLSHRVSGAQKQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 853 EG-LRSSRCNVNHHPIFHEGNGTIVTFDYSKATLDGRMLMRASASSENKASSSKATP 911
QY 901 QLELPVKYAVVMVYVTSHGVSSTKYNF-TASENTSRVMQHOYQVSNLQORSUPISLVFLVP 959
DB 912 QLELPVKYAVVTMISROEESTKYENFATSDCKMKEAHRVYRVNLSQORDLAISINFMPV 971
QY 960 VRLNQTVIWDPPQVTFSENLSSTCHTERLPSSHDSFLAELRKAPVNVNCSTIACQICDI 1019
DB 972 VLLNGVAVMDVVMEMPQSOL--PCVSEKPKPQHSDFLTQISRPMLDCSIADCLQFRCDV 1029
QY 1020 PFFGTQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTE 1079
DB 1030 PSFSVQBELDFTLKGNLISFGWVRETLQKVLVVSVAEITFTDTSVYSQLPGQEAFFRAQME 1089
QY 1080 TKVBPPEVPNPLPIIVGSSVGLLLALITAAALYKLGFFKROYKDMSE 1128
DB 1090 MVLEDEVYNAIPIIMGSSVGALLLALITATLYKLGFFKRYKEMLEB 1138

```

Search completed: November 25, 2003, 14:16:34
Job time : 40.2978 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	5862	99.8	1153	1	US-08-173-497-3	Sequence 3, Appli
2	5862	99.8	1153	1	US-08-286-889-3	Sequence 3, Appli
3	5862	99.8	1153	1	US-08-485-618-3	Sequence 3, Appli
4	5862	99.8	1153	1	US-08-362-652-3	Sequence 3, Appli
5	5862	99.8	1153	2	US-08-605-672-3	Sequence 3, Appli
6	5862	99.8	1153	2	US-08-482-293A-3	Sequence 3, Appli
7	5862	99.8	1153	2	US-08-943-363-3	Sequence 3, Appli
8	5862	99.8	1153	3	US-09-193-043-3	Sequence 3, Appli
9	5862	99.8	1153	4	US-09-688-307A-3	Sequence 3, Appli
10	5831.5	99.2	1152	2	US-08-476-062A-43	Sequence 43, Appl
11	5831.5	99.2	1152	5	PCR-US96-01314-43	Sequence 43, Appl
12	5831.5	99.2	1152	6	5424339-2	Patent No. 5424399
13	3469	59.0	1163	2	US-08-476-062A-44	Sequence 44, Appl
14	3469	59.0	1163	5	PCR-US96-01314-44	Sequence 44, Appl
15	3446	58.6	1163	1	US-08-173-497-4	Sequence 4, Appli
16	3446	58.6	1163	1	US-08-286-889-4	Sequence 4, Appli
17	3446	58.6	1163	1	US-08-485-618-4	Sequence 4, Appli
18	3446	58.6	1163	1	US-08-362-652-4	Sequence 4, Appli
19	3446	58.6	1163	2	US-08-605-672-4	Sequence 4, Appli
20	3446	58.6	1163	2	US-08-482-293A-4	Sequence 4, Appli
21	3446	58.6	1163	2	US-08-943-363-4	Sequence 4, Appli
22	3446	58.6	1163	3	US-09-193-043-4	Sequence 4, Appli
23	3446	58.6	1163	4	US-09-688-307A-4	Sequence 4, Appli
24	3411	58.0	1161	1	US-08-173-497-2	Sequence 2, Appli
25	3411	58.0	1161	1	US-08-286-889-2	Sequence 2, Appli
26	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
27	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli


```
QY 61 RLOQVPEAVNMSLGLSLAATSPPOLLAGCGPTVHQTCSNTYVVKGLCFPGSNLRQOPQK 120
DB 77 RLOQVPEAVNMSLGLSLAATSPPOLLAGCGPTVHQTCSNTYVVKGLCFPGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSSEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSSEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVRVRELFNITNGARKNAFKILVLI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVRVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436
QY 421 QNTGHWESNANVKGTOIGAYFCASICSVDVDSNGSTDVLVIGAPHYEQTGRGQVSVCLP 480
DB 437 QNTGHWESNANVKGTOIGAYFCASICSVDVDSNGSTDVLVIGAPHYEQTGRGQVSVCLP 496
QY 481 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGDVNGDKLTVAITGAPGEEDNRGAVLYF 540
DB 497 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGDVNGDKLTVAITGAPGEEDNRGAVLYF 556
QY 541 HCTSGSGISPSHSQRIAGSKLSPRIQYFQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HCTSGSGISPSHSQRIAGSKLSPRIQYFQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPSPVLRNLF 720
DB 677 YDLALDLSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPSPVLRNLF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLETFALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLETFALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GREFNVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLONORSORSHWLACESASSTEV 840
DB 797 GREFNVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLONORSORSHWLACESASSTEV 856
QY 841 SGALKSTCSINHPIFPENSEVFTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 900
DB 857 SGALKSTCSINHPIFPENSEVFTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNOTVWDRPOVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRTQCDIP 1020
DB 977 RLNOTVWDRPOVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRTQCDIP 1036
QY 1021 FFGIOEFNATLKGNLSDFWYIKTSHNHLIIYSTAILFNDSDVFTLLPQOGAFVRSQTEF 1080
DB 1037 FFGIOEFNATLKGNLSDFWYIKTSHNHLIIYSTAILFNDSDVFTLLPQOGAFVRSQTEF 1096
QY 1081 KVEPFEPVNPPLIVIGSSVGGLLLLALITAAALYKLGFFKRYQKMDMSEGGPPGABPQ 1137
DB 1097 KVEPFEPVNPPLIVIGSSVGGLLLLALITAAALYKLGFFKRYQKMDMSEGGPPGABPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLOTENAMTQENARGQSVVOLQGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCPEPI 60
DB 17 FNLOTENAMTQENARGQSVVOLQGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCPEPI 76
QY 61 RLOQVPEAVNMSLGLSLAATSPPOLLAGCGPTVHQTCSNTYVVKGLCFPGSNLRQOPQK 120
DB 77 RLOQVPEAVNMSLGLSLAATSPPOLLAGCGPTVHQTCSNTYVVKGLCFPGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSSEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSSEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVRVRELFNITNGARKNAFKILVLI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVRVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376
```

QY 361 GVFLYTSKXSTFINMTRVDSMDNDAYLGVAIAAILNRVQSLVGLAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKXSTFINMTRVDSMDNDAYLGVAIAAILNRVQSLVGLAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQSVCPPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQSVCPPL 496
QY 481 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRTAGSKSLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAGQHVLRLRSQ 600
DB 557 HGTSGSGISPSHQSRTAGSKSLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAGQHVLRLRSQ 616
QY 601 PVLVRKAIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
DB 617 PVLVRKAIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
QY 661 YDLALDSGRSHRAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 677 YDLALDSGRSHRAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 736
QY 721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFFPEKNCNDNICODDLSITPFSMLDCLVWG 780
DB 737 SLVGTPLSAFNGLRPVLAEDAQRLLFTALFFPEKNCNDNICODDLSITPFSMLDCLVWG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLACASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLACASSTEV 856
QY 841 SGALKSTSCSINHPFNPENSEVTNITFDVDSKASLGNKLLKANVTSENMPTNKTEF 900
DB 857 SGALKSTSCSINHPFNPENSEVTNITFDVDSKASLGNKLLKANVTSENMPTNKTEF 916
QY 901 QLELPVKIAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYOVSNLQORSPLISLVFLVPV 960
DB 917 QLELPVKIAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYOVSNLQORSPLISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAIPVNCISAVCORIQCDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAIPVNCISAVCORIQCDIP 1036
QY 1021 PFGIQEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFVRSQTET 1080
DB 1037 PFGIQEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFVRSQTET 1096
QY 1081 KVEPFEVNPPLIVGSSVGLLLALITAAALKYKGFQYKDMWSEGGPGCAEPQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGLLLALITAAALKYKGFQYKDMWSEGGPGCAEPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCE: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun.
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQCSTENTYVKGICFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTLLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPISTVMEOLKSKTLLFSLMOYSEEF 196
QY 181 RIHFTFKEFQNNPNRSLIKPITQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTFKEFQNNPNRSLIKPITQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVDGDAFRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKXSTFINMTRVDSMDNDAYLGVAIAAILNRVQSLVGLAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKXSTFINMTRVDSMDNDAYLGVAIAAILNRVQSLVGLAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQSVCPPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQSVCPPL 496
QY 481 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRTAGSKSLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAGQHVLRLRSQ 600

557 HCTSGGSPSHSQRASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 616
601 PVLRYKAIEMFNPREVARNVFNCDQVYVKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLRYKAIEMFNPREVARNVFNCDQVYVKEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGRPHSRVAVFNETKSTRQTQVGLTQTCETLKLQPNCEIDPSPVLRNLF 720
677 YDLALDSGRPHSRVAVFNETKSTRQTQVGLTQTCETLKLQPNCEIDPSPVLRNLF 736
721 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
797 GPREFNVTVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
841 SCALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
857 SCALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 960
917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 976
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1020
977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCIP 1036
1021 PFGIOBEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTE 1080
1037 PFGIOBEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQAFVRSQTE 1096
1081 KYPEFVNPPLIIVGSSVGGLLALITAAALYKLGFFKRYQKDWMSGGPPGABPQ 1137
1097 KYPEFVNPPLIIVGSSVGGLLALITAAALYKLGFFKRYQKDWMSGGPPGABPQ 1153

RESULT 4
US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLOQSRVVGAPQEIIVANQORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOQSRVVGAPQEIIVANQORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMVSLGSLAATTPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQOPQK 120
DB 77 RLQVPEAVNMVSLGSLAATTPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQOPQK 136
QY 121 PPEALRCQPEDSDIAFLVDGSGSIIIPHDFRRAKEPISTVMEQKKKSLFSLMOYSEEP 180
DB 137 PPEALRCQPEDSDIAFLVDGSGSIIIPHDFRRAKEPISTVMEQKKKSLFSLMOYSEEP 196
QY 181 RIHTTFEFQNNPNSLIKPIITOLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240
DB 197 RIHTTFEFQNNPNSLIKPIITOLLGRTHATGIRKVVRELFINITGARKNAFKILIVI 256
QY 241 TDGKFKDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDRHVQIN 300
DB 257 TDGKFKDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDRHVQIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPSEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSPSEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
DB 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 600
DB 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 616
QY 601 PVLRYKAIEMFNPREVARNVFNCDQVYVKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRYKAIEMFNPREVARNVFNCDQVYVKEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAVFNETKSTRQTQVGLTQTCETLKLQPNCEIDPSPVLRNLF 720
DB 677 YDLALDSGRPHSRVAVFNETKSTRQTQVGLTQTCETLKLQPNCEIDPSPVLRNLF 736
QY 721 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNGRLPVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
Qy 1021 FFGIOEFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSQDET 1080
Db 1037 FFGIOEFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSQDET 1096
Qy 1081 KVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKXDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKXDMMSSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQOPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVYKGLCFGLFGLSRLQOPQK 120
Db 77 RLQOPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVYKGLCFGLFGLSRLQOPQK 136
Qy 121 FPEALRGCCQEDSDIAFLVDGSGSIIPHPDRAKEPISTVMEOLKSKTKTLFSLMQYSSEF 180
Db 137 FPEALRGCCQEDSDIAFLVDGSGSIIPHPDRAKEPISTVMEOLKSKTKTLFSLMQYSSEF 196
Qy 181 RIHFTPEKFNQNPNSRLIKPIITQLLGRTHRTATGIRKVVRELPIININGARKNAFKILILI 240
Db 197 RIHFTPEKFNQNPNSRLVKPIITQLLGRTHRTATGIRKVVRELPIININGARKNAFKILVVI 256
Qy 241 TDGEKEGDPGLGYEDVPEADREGVIRYVIGVDGAPSEKSRQELNTVASKPRDHPVQIN 300
Db 257 TDGEKEGDPGLGYEDVPEADREGVIRYVIGVDGAPSEKSRQELNTVASKPRDHPVQIN 316
Qy 301 NFEALXTIQNOLREKIFAIEGTQSGSSSEFHEMSOEGFSAATISNGPLLLSTVGSYDWAG 360
Db 317 NFEALXTIQNOLREKIFAIEGTQSGSSSEFHEMSOEGFSAATISNGPLLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGAYAAAILLRNVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNAYLGAYAAAILLRNVQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQSVCP 496
Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVDLGDVNGDKLTDAIICAPGEEDNRGAYVLF 540
Db 497 PRGORARWOCDAVLYGEOGQPMGRFGAALTVDLGDVNGDKLTDAIICAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTGGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTGGAQGHVLLRSQ 616
Qy 601 PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDRLEQIQSVVT 660
Db 617 PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDRLEQIQSVVT 676
Qy 661 YDLALDSGRPHSRVFNKSTRTOTQVLGLTQTCETLKLQLPNCIEDPVSPILVRLNF 720
Db 677 YDLALDSGRPHSRVFNKSTRTOTQVLGLTQTCETLKLQLPNCIEDPVSPILVRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAORLFTALPPFEKNKCGNDNICQDDLSITTFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAORLFTALPPFEKNKCGNDNICQDDLSITTFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036

Qy	1021	FFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET	1080
Db	1037	FFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET	1096
Qy	1081	KVEPFEVNPPLPLTVGSSVGLLALITAAALYKLGFFKQYKDMSEGGPPGAEPO	1137
Db	1097	KVEPFEVNPPLPLTVGSSVGLLALITAAALYKLGFFKQYKDMSEGGPPGAEPO	1153
RESULT 6			
US-08-482-293A-3			
; Sequence 3, Application US/08482293A			
; Patent No. 5831029			
; GENERAL INFORMATION:			
; APPLICANT: Gallatin, W. Michael			
; APPLICANT: Van der Vieren, Monica			
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit			
; NUMBER OF SEQUENCES: 103			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun			
; STREET: 233 South Wacker Drive, 6300 Sear Tower			
; CITY: Chicago			
; STATE: Illinois			
; COUNTRY: United States			
; ZIP: 60606-6402			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/482,293A			
; FILING DATE:			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/173,497			
; FILING DATE: 23-DEC-1993			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/286,889			
; FILING DATE: 5-AUG-1994			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/362,652			
; FILING DATE: 21-DEC-1994			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Williams Jr., Joseph A.			
; REGISTRATION NUMBER: 38,659			
; REFERENCE/DOCKET NUMBER: 27866/32684			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 312-474-6300			
; TELEFAX: 312-474-0448			
; TELEX: 25-3856			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1153 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-482-293A-3			
Query Match 99.8%; Score 5862; DB 2; Length 1153;			
Best Local Similarity 99.3%; Pred. No. 0;			
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ENLDTENAMTFQENARGFGQSVVQLOGSRVVVQCAPQEIIVAAORGSLYQCDYSTGSCPEI	60
Db	17	ENLDTENAMTFQENARGFGQSVVQLOGSRVVVQCAPQEIIVAAORGSLYQCDYSTGSCPEI	76
Qy	61	RLQVPVAVNMVSLGLSLAATTPSPQLIACGPTVHQTCSENTYVYKGLCFLFGSNLRQOPQK	120
Db	77	RLQVPVAVNMVSLGLSLAATTPSPQLIACGPTVHQTCSENTYVYKGLCFLFGSNLRQOPQK	136
Qy	121	FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKXSKTLFSLMQYSEEF	180

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

```

; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 127866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-363-3

Query Match          99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY      1  FNLDTENAMTFQBNARGFGQSVVQLQGRVWVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB      17  FNLDTENAMTFQBNARGFGQSVVQLQGRVWVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY      61  RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 120
DB      77  RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 136

QY      121  FPEALRGCEQSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKSKTFLSLMQVSEEF 180
DB      137  FPEALRGCEQSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKSKTFLSLMQVSEEF 196

QY      181  RIHFTFKFQNNPNPSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKLLI 240
DB      197  RIHFTFKFQNNPNPSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKLLI 256

QY      241  TDGEKFGDPLGVEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFQIN 300
DB      257  TDGEKFGDPLGVEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFQIN 316

QY      301  NFEALKTIONQREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG 360
DB      317  NFEALKTIONQREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG 376

```

RESULT 8

US-09-193-043-3

; Sequence 3, Application US/09193043

; Patent No. 6251395

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

; APPLICANT: Van der Vieren, Monica

; TITLE OF INVENTION: No. 6251395el Human 2

; FILE REFERENCE: 27866/35004

; CURRENT APPLICATION NUMBER: US/09/193,043

; EARLIER FILING DATE: 1998-11-16

; EARLIER FILING DATE: 1993-12-23

; EARLIER FILING DATE: 1994-08-05

; EARLIER FILING DATE: 1994-08-05

; EARLIER FILING DATE: 1994-12-21

; EARLIER FILING DATE: 1997-10-03

```

QY      361  GVFLVTSKEKSTFFINMTRVDSMDNDAYLGYAAAAIIILNRVQSLVILGAPRYQHIGLVAMFR 420
DB      377  GVFLVTSKEKSTFFINMTRVDSMDNDAYLGYAAAAIIILNRVQSLVILGAPRYQHIGLVAMFR 436

QY      421  QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSDTLVLIGAPHYVYQTRGGQVSVCP 480
DB      437  QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSDTLVLIGAPHYVYQTRGGQVSVCP 496

QY      481  PRGQARWQCDVAVLYGEOQPMGRFGAALTVLVDVNGDKLTDAVTCAPCEEDNRGAVVLF 540
DB      497  PRGQARWQCDVAVLYGEOQPMGRFGAALTVLVDVNGDKLTDAVTCAPCEEDNRGAVVLF 556

QY      541  HGTSGSISPSHSORIASKJLSPRLQYFGQSLSGQDLTMDGLVDLTVAQAGHVLLLRQ 600
DB      557  HGTSGSISPSHSORIASKJLSPRLQYFGQSLSGQDLTMDGLVDLTVAQAGHVLLLRQ 616

QY      601  PVLVRKAIAMEFNPREVARNVFECDNVVYKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 660
DB      617  PVLVRKAIAMEFNPREVARNVFECDNVVYKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 676

QY      661  YDLALDSGRPHSRVAFNETKNSSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
DB      677  YDLALDSGRPHSRVAFNETKNSSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736

QY      721  SLVGTPLSAFGLNLRPVLAEQAQLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
DB      737  SLVGTPLSAFGLNLRPVLAEQAQLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796

QY      781  GPREFNVTVVRNDCEDSYRQVTFPFLDLISYRKVSTLQNRORSORSWRLACESASTEV 840
DB      797  GPREFNVTVVRNDCEDSYRQVTFPFLDLISYRKVSTLQNRORSORSWRLACESASTEV 856

QY      841  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
DB      857  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916

QY      901  QLELPKYAVVMVVTSHGVSTKYLNFTASENTRVMQHQYQVSNLQGRSLPISLVFLVPV 960
DB      917  QLELPKYAVVMVVTSHGVSTKYLNFTASENTRVMQHQYQVSNLQGRSLPISLVFLVPV 976

QY      961  RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB      977  RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036

QY      1021  FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEIILFNDSVFTLLPQCGAFVRSOTET 1080
DB      1037  FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEIILFNDSVFTLLPQCGAFVRSOTET 1096

QY      1081  KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYOKDMWSEGGPPGABPQ 1137
DB      1097  KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYOKDMWSEGGPPGABPQ 1153

```


; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARFGOSVWQVLOGSRVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGOSVWQVLOGSRVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILIVI 256
Qy 241 TDGKFGDPLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTIASKPPRDHVFQIN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOCHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOCHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYTEQTRGGQVSCPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTEQTRGGQVSCPL 496
Qy 481 PRGQRRQWQDAVLGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQRRQWQDAVLGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVT 660
Db 617 PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVT 676
Qy 661 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGMLRPLVLAEDAQRLLTALPPEKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGMLRPLVLAEDAQRLLTALPPEKNGCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GREFENVTVVRNDCGDSYRTQVTFPPFLDLRYKVTSLQNSORSWRLACESASSTEV 840
Db 797 GREFENVTVVRNDCGDSYRTQVTFPPFLDLRYKVTSLQNSORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 991 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVWQHOYQVSNLQORSLPISLFLVPV 960

Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVWQHOYQVSNLQORSLPISLFLVPV 976
Qy 961 RLNQTVINDRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVNCSTAVCORIQCDIP 1020
Db 977 RLNQTVINDRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVNCSTAVCORIQCDIP 1036
Qy 1021 FFIQIEEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQOGAFVRSQTET 1080
Db 1037 FFIQIEEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQOGAFVRSQTET 1096
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPO 1137
Db 1097 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSGEGPPGAEPO 1153

RESULT 9

US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARFGOSVWQVLOGSRVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGOSVWQVLOGSRVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILIVI 256
Qy 241 TDGKFGDPLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTIASKPPRDHVFQIN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCVPL 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCVPL 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLRQVFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLRQVFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 616
Qy 601 PVLRYKAIMENPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRYKAIMENPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPSRAVENETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPSRAVENETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNIQDDLSITFFSMSLDCLVWG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNIQDDLSITFFSMSLDCLVWG 796
Qy 781 GPREFNVTVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 976
Qy 961 RLNTQVIWDRPOVTFSENLSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDIP 1020
Db 977 RLNTQVIWDRPOVTFSENLSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDIP 1036
Qy 1021 FFGIOEEFNATLKGNSLSDWIKTSHNHLIIVSTAELFNDSVETLLPGQAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNSLSDWIKTSHNHLIIVSTAELFNDSVETLLPGQAFVRSOTET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGAEPO 1137
Db 1097 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGAEPO 1153

RESULT 10

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275

GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476, 062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216, 081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637, 830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539, 842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212, 573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29, 066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSAAATTPOLLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSAAATTPOLLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 136
Qy 121 FPEALRCPCOEDSDIAFLIDSGSGIIIPHDRFRAKEFISTVMEQLKKSKTLFSLMQYSEEF 180
Db 137 FPEALRCPCOEDSDIAFLIDSGSGIIIPHDRFRAKEFISTVMEQLKKSKTLFSLMQYSEEF 196
Qy 181 RIHPTFKFQNNPNRSLRKPITQLLORTHTATGIRKVRRELFNITNGARKNAPKILVLI 240
Db 197 RIHPTFKFQNNPNRSLRKPITQLLORTHTATGIRKVRRELFNITNGARKNAPKILVLI 256
Qy 241 TDCEKFGDPLGYEDVPEADREGVIRVYVIGVDAFRSEKSRQELNITVASKPPRDHVFQIN 300
Db 257 TDCEKFGDPLGYEDVPEADREGVIRVYVIGVDAFRSEKSRQELNITVASKPPRDHVFQIN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCVPL 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQVSCVPL 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLRQVFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 600

556 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 615
601 PVLRVKAIMFNPVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
616 PVLRVKAIMFNPVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
661 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQPNCEIDPVPVLRNLF 720
676 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQPNCEIDPVPVLRNLF 735
721 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFEKNCNDNI CQDLSITFSFMSLCLVVG 780
736 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFEKNCNDNI CQDLSITFSFMSLCLVVG 795
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
841 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
856 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
901 QLELPVKYAVYVMTSHGVSTKYNFTASNTSRVVMQHOYQVSNLQQRSLPISLVFLVPV 960
916 QLELPVKYAVYVMTSHGVSTKYNFTASNTSRVVMQHOYQVSNLQQRSPPIISLVFLVPV 975
961 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
976 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
1021 FPGIOBEFNATLKNLSFDWYKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080
1036 FPGIOBEFNATLKNLSFDWYKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1095
1081 KVEPFEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1137
1096 KVEPFEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1152

RESULT 11

PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29, 066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906

TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.2%; Score 5831.5; DB 5; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSYQCDYSTGSCSEPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSYQCDYSTGSCSEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCENTYVKGCLFGLSGNLRQOPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCENTYVKGCLFGLSGNLRQOPOK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPRRAKEFIPTVMEQKKSKTFLSLMOYSBEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPRRAKEFIPTVMEQKKSKTFLSLMOYSBEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLGRTHATGIRKVVRELFINITNGARKNAFKILLI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLGRTHATGIRKVVRELFINITNGARKNAFKILLI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 316
QY 301 NFEALKTQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWMAG 360
DB 317 NFEALKTQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWMAG 376
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEGTREGQVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEGTREGQVSVCLP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTVLGDVNGDKLTDAI GACPGEENRGAVALF 540
DB 497 PRG-RARWOCDAVLVGEQOPWGRFGAALTVLGDVNGDKLTDAI GACPGEENRGAVALF 555
QY 541 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 556 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 615
QY 601 PVLRVKAIMFNPVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 616 PVLRVKAIMFNPVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQPNCEIDPVPVLRNLF 720
DB 676 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQPNCEIDPVPVLRNLF 735
QY 721 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFEKNCNDNI CQDLSITFSFMSLCLVVG 780
DB 736 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFEKNCNDNI CQDLSITFSFMSLCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTFTFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915

Qy 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 975
Qy 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1020
Db 976 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1035
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIOBEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALLITAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLLALLITAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78.871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539.842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212.573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 1152
5424399-2

Query Match 99.28; Score 5831.5; DB 6; Length 1152;
Best Local Similarity 99.08; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIIVAAANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKVLGCLFLGSLNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKVLGCLFLGSLNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPDFRRAKEFISTVMEQLKSKTILFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPDFRRAKEFISTVMEQLKSKTILFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPRSLIPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPRSLIPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILIVI 256
Qy 241 TDGKFGDPLGYEDVIEADREGVIRYVIGVDGDAFRSEKSRQELNVTASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVIEADREGVIRYVIGVDGDAFRSEKSRQELNVTASKPPRDHVFQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQEGFSAATISNGPLLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQEGFSAATISNGPLLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTFRVDSMDNDAYLGVAALILNRRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTFRVDSMDNDAYLGVAALILNRRVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVSCPL 496
Qy 481 PRGQARWQCDAYLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRRGAYVLF 540

Db 497 PRG-RARWQCDAYLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRRGAYVLF 555
Qy 541 HGTSGGIGSPSHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
Db 556 HGTSGGIGSPSHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 615
Qy 601 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
Db 616 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 675
Qy 661 YDLALDSGPHSRVAVNETKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 YDLALDSGPHSRVAVNETKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 735
Qy 721 SLVGTPLSAFNGNLRPVLAEDAQRLFTALFPFENKCGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGNLRPVLAEDAQRLFTALFPFENKCGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 796 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKXANTVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKXANTVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 975
Qy 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1020
Db 976 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1035
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIOBEFNATLKNLSFDWYIKTSHNHLITVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALLITAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLLALLITAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 13
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

/	FILING DATE:	18-JUN-1990
/	APPLICATION NUMBER:	07/212,573
/	FILING DATE:	28-JUN-1988
/	ATTORNEY/AGENT INFORMATION:	
/	NAME:	Freeman, John W.
/	REGISTRATION NUMBER:	29,066
/	REFERENCE/DOCKET NUMBER:	100786/068003
/	TELECOMMUNICATION INFORMATION:	
/	TELEPHONE:	617/542-5070
/	TELEFAX:	617/542-8906
/	TELEX:	200154
/	INFORMATION FOR SEQ ID NO:	44:
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH:	1163 amino acids
/	TYPE:	amino acid
/	TOPOLOGY:	linear
/	MOLECULE TYPE:	protein
/	US-08-476-062A-44	

		Query Match	59.0%; Score 3469; DB 2; Length 1163;
		Best Local Similarity	61.0%; Pred. No. 7.le-284;
		Matches	689; Conservative 141; Mismatches 293; Indels 6; Gaps 6;

Qy	1	FNLDTENAMTFQENARGFGQSUVLOQSRVWVGAPQEIIVAAQNQRGSLYQCDYSTGSCEPI	60
Dd	20	FNLDTBELTAFRVDSAGFGSDVVQYANSVWVGAPQKITAAONTGGLYQCGYSTGACEPI	79
Qy	61	RLOVPVEAVNMSLGLSLAAATSPPOLACGPTVHOTCSENTYVKGLCFLFQSNLRQQPQK	120
Dd	80	GLOVPPEAVNMSLGLSLASTSPSQLLACGPTVHHEGRNMYLTGLCLFLGPT--QLTQR	137
Qy	121	FPBALRGCPQEDSDIAFLVDGSGSIIPHFERRAKEFISTMVQELKKSKTLFSLMQYSSEF	180
Dd	138	LPVSREQCPREQDIVEIDGSGSISRRNFATMNFVRAVISQFORSTQFSLMQFSNK	197
Qy	181	RIHTTKPFQNNPNPRSLIPIITOLLGRTHATGIRKVVRRLFNITWGAKNAFKILLI	240
Dd	198	QTHFTFEERFTNSPLLASVHQLOGFTYTATAIQNVVHLRPHASYGARRDATKLIVI	257
Qy	241	TDEKEGDPLGYEDVIPADREGVIRVYGVGDAPREKSROELNTVASKPPRDHVFOIN	300
Dd	258	TDGKGSGSLDYKDVI PMADAAGII RYAIGVLAFQNRNSWKELNDIASKPSQEHIFKVE	317
Qy	301	NFEALKTIONQLREKIPAIBCTQTGSSSSFEHMSQBGFSAAITNSGPLLSTVGSYDWAG	360
Dd	318	DFDALNDIONQLKEKIPAIEGTETTSSSSELENMAQEGSFVFTPDGPVLGA VG SFTWSG	377
Qy	361	GVPLYTSKESKSTFINMTTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPPRYOHIGLUAMFR	420
Dd	378	GAFLYPPNMPTFINMSQENVDMRDSYLGYSTELALWKGVOSLVLGAPRYOHTKAVIFT	437
Qy	421	QNTGMWESNANVKGTOICAYFGASLCSDVDNSGSTDLVLIGAPHYTEOTRGGOVSVCP	480
Dd	438	QVSRQRMKAEEVTGTQISYFGASLCSDVDVTDGSTDVLVLIGAPHYTEQTRGGQVSVCP	497
Qy	481	PRGORARWCDAVLYGFGQGPWGRFGAALT VLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Dd	498	PRGWR-RWWCDAVLYGEGHPGWGRFGAALT VLGDVNGDKLTDVIGAPGEENRGAVYLF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPLRYQFQSLGSGQDLTMDBGLVDLT TVGAQHVL LRSQ	600
Dd	557	HGVLGPSISPSHSQRIAGSKLSRLRYQFQALSGQDLTDGBLDVLA VARGQV LLLRTR	616
Qy	601	PVLRVKAIMERNPREVARNYPECNDVVVKGEAGEVRVCLHVQKSTRDRLRREGIOQSVVT	660
Dd	617	PVLWMVGSMQPIPAEIRSAFRECQVWSEQTLVQSNICLYDKRKSNLLGSRDLQSSVT	676
Qy	661	YDALDSCRPHSRVAFNETKNSTRQTQVLGLTQTCETLKQLQNPNCIEDPVSPVILRLNF	720
Dd	677	LDALDPGRSLSPRATFQETKNRSISRVRVLGKAHCENFNLLLPCVEDSVTPITRLNF	736
Qy	721	SLVGTPLPSAFGNLRPVLAEDAQRFLTALFPFEKNCNGNDNI CODDSLITFFSMSLDCLVVG	780

Tue Nov 25 14:52:01 2003

Db 318 DFDALKIQNOLKEKIFAIBGTETETISSSSFELEMAQEGFSAVFTPDGPVLGAVCSFTWSG 377
QY 361 GVFLYTSKEKSTFINMTFRVDSMDMDAYILGYAAAIILRRNRVQSLVGLGAPRYOHIGLVAMER 420
Db 378 CAFLYPPNMSFTFINMSQENVDMRDSVILGYSTELALWKGVQSLVGLGAPRYOHICKAVIFI 437
QY 421 QNTGMEANVAKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYETQTRGGQVSVCP 497
QY 481 PRGQARWQCDVILYGEQGPWKREGAALTVLGDVNGDKLTDVAIGAPGEBEDNRGAVYLF 540
Db 498 PRGMR-RWMCDAVLYGEGHPWGRFGAALTVLGDVNGDKLTDVIGAPGEBEENRGAVYLF 556
QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIEMFNPREVANRPECNDQVVKKEAGEVRVCLHVQKSTRDRRLREGQIQSVTT 660
Db 617 PVLWVGVSQFIPASIPRSAFEPCREQVYSEQTLVQSNICLYDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVFNENKSTRTQTVLGLTQTCETLKLQPCNIEBDVPSPVILRLNF 720
Db 677 LDALAPGRLSPRAFQETKNRSLSRVRLGKAHCENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAFGLRLPVLAEADAQRLFTALPPEKNCNGDNICODDLSITFSFMSLCLVVG 780
Db 737 TLVGPLLAFLRLPMLAALQRYFTASLPFEKNCGADHICODNLGIFSFPGLKSLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPFLDLSYRKVSTLQONORSORSWRLACESASSTEV 840
Db 797 SNLELNAEVMVMNDGEDSVGTTITFHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANVTSENMMPTNKTEF 900
Db 855 SQGTWSTSCRINHLPFRGGAQITFLATFDVSPRAVGLDRLLLIANVSSENNIPRISKTI 914
QY 901 QLELPVKYAVYVVTSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLQORSPLISLVFLVP 959
Db 915 QLELPVKYAVYVVSHEQFTKYNFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPOVNCNIAVCORIQCDI 1019
Db 975 VELNQEAVMMDVEVSHFQNPNSLRCSSEKIAPPASDFLAHQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFTGQEEFNATLKGNLSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPCQGAFAVRSOTE 1079
Db 1035 PSFSVQEEFDTLLKGNLSFGWVRQIILQKKVSVVSVAEIIPDTSVVSQLPQGEAFMRAQTI 1094
QY 1080 TKVEPPEVPNPLPIVSSVGGLLLALITAALYKLGFFKRYQKDMWSE 1128
Db 1095 TVLEKYKVNPIELIVGSSIGGLLLALITAVLYKVGFYKRYKEMWEE 1143

Search completed: November 25, 2003, 14:23:21
Job time : 17.5618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGPGPABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	1 RWU1B	cell surface glyco
2	4476	76.2	1153	2 S00551	leukocyte surface
3	3483	59.3	1163	1 RWU1C	cell surface glyco
4	1548.5	26.4	1170	2 S03308	cell surface glyco
5	1533.5	26.1	1163	2 I56126	lymphocyte fuction
6	1149	19.6	1179	2 A53213	integrin alpha-E c
7	1102.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1084	18.5	1170	2 I45914	integrin alpha 2 s
9	1072	18.2	1178	2 S44142	VLA-2 protein homo
10	1069	18.2	1181	2 A33998	integrin alpha-2 c
11	1060	18.0	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	633	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha-2 c
16	579.5	9.9	1054	2 JCY7294	alphanp integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	543.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 B27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	530.5	9.0	1072	2 A38457	integrin alpha-6 c
25	529.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.5	1146	2 S40311	integrin - fruit f

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,

B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:CROSS-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148

A>Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A:CROSS-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA5

A>Note: The authors translated the codon TAC for residue 1129 as Thr

A>Note: part of this sequence, including the amino end of the mature protein, was confi

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally i

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:CROSS-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A:Reference number: A34193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:CROSS-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:CROSS-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FLE>
A:CROSS-references: GB:S52227; NID:9263047; PIDN:AB24821.1; PID:9263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PIE>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A>Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:CROSS-references: GB:M84477; NID:gl80184; PIDN:AAA51960.1; PID:g553219
A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:CROSS-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VMA2>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.9% Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDSYSCPEI 60
Db 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDSYSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCFPGSLNRQQPK 120
Db 77 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCFPGSLNRQQPK 136

Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSVVMEQLKSKTLFSLMOYSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSVVMEQLKSKTLFSLMOYSEF 196

Qy 181 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIIRLLNITNGARKNAFKILVI 240
Db 197 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIIRLLNITNGARKNAFKILVI 256

Qy 241 TDGEKFGDPLGVEDYVPEADREGVIRYVIGVDGAFRSKSKQLNTIASKPPRDHFQVN 300
Db 257 TDGEKFGDPLGVEDYVPEADREGVIRYVIGVDGAFRSKSKQLNTIASKPPRDHFQVN 316

Qy 301 NFELAKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATITNGPLLLSTVGSYDWAG 360
Db 317 NFELAKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATITNGPLLLSTVGSYDWAG 376

Qy 361 GVFLYTSKESKSTFINNTRVDSNDMDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINNTRVDSNDMDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 496

Qy 481 PRGORARWQCDAYLYGEQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDAYLYGEQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616

Qy 601 PVLVRKAIMEFNPREVARNVFECDVVGKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECDVVGKGEAGEVRVCLHVQKSTRDRREGIOISVVT 676

Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVRLNF 736

Qy 721 SLVGTPLSAFNGLRPVLAEDAQRALTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEDAQRALTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKYSTLQNRQSRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKYSTLQNRQSRWSRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

Qy 901 QLELPVYAYVMVVTSHGVSTKYLNTFASENTSRVMQHQVQVSNLQSRSLPISLVFLVPV 960
Db 917 QLELPVYAYVMVVTSHGVSTKYLNTFASENTSRVMQHQVQVSNLQSRSLPISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP 1036

Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1096

Qy 1081 KVPEFEPNPLPLIVGSSVGLLLALITAAALYKLGFFKRYKQKDMWSEGGPPGAEPO 1137
Db 1097 KVPEFEPNPLPLIVGSSVGLLLALITAAALYKLGFFKRYKQKDMWSEGGPPGAEPO 1153

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C:Accession: S00551; I59078

R:Pytela, R.

EMBO J. 7, 1371-1378, 1988

A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th

A:Reference number: S00551; MUID:88312584; PMID:3044779

A:Accession: S00551

A:Molecule type: DNA

A:Residues: 1-1153 <PYT>

A:CROSS-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983

A>Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Toplew, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:G198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

	Query Match	76.2%;	Score 4476;	DB 2;	Length 1153;
	Best Local Similarity	74.3%;	Pred. No. 1.2e-301;		
	Matches 845;	Conservative 142;	Mismatches 149;	Indels 2;	Gaps 2;
Qy	1	FNLDTENAMTFOENARGFGOSVVOLOGSRVVVGAPQEIIVAANORGSLGYQCDYGTGSCPEI	60		
Db	17	FNLDETHPMTFQENAKFGQNVVLOGGTSVVVAAPQAKAVQTGALYQCDYGTSTRCHPI	76		
Qy	61	RLOVPVPAVNMSLGLSLAATTSPQLLACAGPTVHQTCSNTYVKGCLCFGLFGSNLRQOPOK	120		
Db	77	PLQVPPPAVNMSLGLSLAVSTVPQQLLACAGPTVHQNCKENTYNGLCYLFGLSNLRPPQO	136		
Qy	121	FPEALRCPOEDSDIAFLDGSIGIIIPHDPRMKKEFVSTVMEQLKSKTTLFSLMQVSEEF	180		
Db	137	FPEALRCPOQESDIVEPLDGSIGINNIDFQMKKEFVSTVMEQPKSKTTLFSLMQYSDEF	196		
Qy	181	RIHPTFKFQNNPNRSLVAKPIITOLLGRTHATGVRKVIARELLNITNGAKNAFKILIVI	240		
Db	197	RIHPTFNDFKRNPSPRHVSPIKOLNGRTKTASGIRKVRKELFHTKNGAKENAAKILVVI	256		
Qy	241	TDGEKFGDPLGYEDVIEADREGVIRVIVIGVDAFRSEKSRQELNTIATSPRHDVFEQVN	300		
Db	257	TDGEKFGDPLDYKDVIPEADRAGVIRVIVGVGNAPKPSRRELDITIASKPAGEHVQVD	316		
Qy	301	NFEALKTIQNOLREKPIFAEGTQTGSSSSSEHEMSQSGFSAATSNAGPLISTVGSYDNAG	360		
Db	317	NFEALNTIQNOLQEKIFAIEGTQTGTSSTSEHEMSQSGFSASITNSGPLSGVGSFDWAG	376		
Qy	361	GVFLYTSKESKSTFTNMTRVDSMDNDAYLGAAAIILNRRVQSLVGLAPRQCHTGLVAMFR	420		
Db	377	GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAILNRRVQSLVGLAPRYQHIGLVVMFR	436		
Qy	421	QNTGMWSNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCP	480		
Db	437	ENFGTWEPHTSIKSGQIGSFYFGASLCSVDMDADGNTNLILIGAPHYYEKTTRGGQVSVCP	496		
Qy	481	PRGORARWQCDVILYGBQGPWGRFGAALTVLGVDPNGDKLTDVAIGAPGBEDNRRGAYILF	540		
Db	497	PRG-RARWQCEALLHGDQHPWGRFGAALTVLGVDPNGDKLTDVAIGAPGBEQNGQAVYIF	555		
Qy	541	HGTSGSGISPSHSORIAGSKLSPLQYFGOSLGGODLTMDGLVDLTVGAGQHVLLLRQ	600		
Db	556	YGASIASLSASHSHRIICAHFSPLGYFGOSLGGKDLTMDGLMDLAVGAQGHLLLRQAQ	615		
Qy	601	PVLRVKAIMEPNPREVARNVPECNQDVVVKGEACEVRVCLHVQKSTRDLREGOIQSVWT	660		
Db	616	PVLREATMEFSPKKVARSVPACQEVQVKNKADGEVRVCLVRKNTKDRLREGDIOSTVT	675		
Qy	661	YDLALDSGRPHSRVAFVNETKNSTRQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF	720		
Db	676	YDLALDPVRSIRIRAFFDETQNTRRRTQVFLGMQKCEITKLILPDCVDDSVSPILRLNY	735		
Qy	721	SLVGTPLISAFGNLRPVLAEDAQRLFTALFPPEKKNCGNDNICODDLSTIFFSPMSLDCLVWG	780		
Db	736	TLVGEPLRSFNGLRPVLAEDAQRFETAMFPPEKKNCGNDNICODDLSTITMGAMGLDTLVWG	795		

Query Match	59.3%;	Score 3483;	DB 1;	Length 1163;
Best Local Similarity	61.4%;	Pred. No. 8.e-233;		
Matches	693;	Conservative	138;	Mismatches 292; Indels 6; Gaps 4;
Qy	1	FNLDTENAMTFOENARGFGSVVLOGSRVVGAPQOEIVAAQNGSLYQCDYSTGSCRP	60	
Db	20	FNLDTBELTAFRVDAGFSDVVQYANSWVVGAPQKITAAQNTGGLYQCYSTGACPE	79	
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQCSTENTYVKGCLFLGSLNRQOPOK	120	
Db	80	GLQVPPVAVNMSLGLSLAATTSPPQLLACGPTVHQCSTENTYVKGCLFLGSLNRQOPOK	137	
Qy	121	FPEALRCQOEDSDIAFLIDGSGSIIPHDFRRKKEFVSTVMQOLKSKTLFSLMQVSEEF	180	
Db	138	LPVSRQECPRQEQDQIVFLIDGSGSISRNFATMNFVRAVISQFQRPSTQFSLMQVSNKF	197	
Qy	181	RIHFTFKFQNNPNSRLKPIITOLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	240	
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGTFTTATQNVHRLFHASYGARRATKILIVI	257	
Qy	241	TGEEKFGDPLGYEDVIEADREGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVPQVN	300	
Db	258	TGCKEGSDSLDYKDVIPMAAAGIIRYAIGVGLAFQNRNSWKELNIDASKPSQEHIFKVE	317	
Qy	301	NFEALKTIQNLREKIPAEITGOTGSSSSFEHMSQEGFSAITSTNGPLLSTVSGYDWAG	360	
Db	318	DFDALKDIONLKEKIFAIEGTETSSSFELEMAQEGFSAVFTPDGVLGAVGVSFTWSG	377	
Qy	361	GVFLYTSKEXSTFTNMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHLGLVAMFR	420	
Db	378	GAFLYPPNMSPTFINMSQENVMDRSDYLGSTELALWKGVQSLVILGAPRYOHTGKAVIFT	437	
Qy	421	QNTQWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEOTRGGQVSCPL	480	
Db	438	QVSRQWRMAKAVETGTQIGSYFGASLCSVDVDTGSDTLVLGAPHYYEOTRGGQVSCPL	497	
Qy	481	PRGORARWQDAVLYGQGGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF	540	
Db	498	PRGWR - RWCWDVLYGQGGQWGRFGAALTVLGVDVNGDKLTDVVGAPGEENRGAVYLF	556	
Qy	541	HGTSGSGISPSHSQRTAGSKLSPLQYFGQSLGGQDLTMDGLAVDLTVGAGQHVLRLSQ	600	
Db	557	HGVLGSPISPSHSQRTAGSKLSRLQYFGQALSGQDLTQDGLVLDVAVGARGQVLLRLTR	616	
Qy	601	PVLRYKAIMENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTDRDLREGQIQSVVT	660	
Db	617	PVLVWGVSMQFIPAEIPRSFAFEQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676	
Qy	661	YDLALDSRPHSRVAFNETKNSRROTQVGLGLTQTCETLKLQLENCIEDPVSPVILRLNF	720	
Db	677	LDLALDPGLRSLPRATFQETKNRSLSRVRVLGLKAHCENFNLLPSCVEDSVPTILRLNF	736	
Qy	721	SLVGTPLSFAFGNLRPVLAEDAQRFLFTALFPPEKNCNGNDNIQDDLSITTSFMSLDCLVWG	780	
Db	737	TLVGKPLAFNLRPLMAADAQRYFTASLPPEKNCAGADHIQDNLGHSISFPGKLSLLVG	796	
Qy	781	GPREFVTVTVRNDGEDSYRTQVTFPPFLDLSYKRVKSTLQNRQSRQSWRLACESASSTEV	840	
Db	797	SNLELNAEVMVNDGEDSYGTTITFSHPAGLSYRVYAEQGQQLSLHLTCDSPAGV--	854	
Qy	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET	900	
Db	855	SGQWTSTSCRINHILIFGGAQITFLATFDVSPRAVLGRDLTLTANVSSENTPRTSKTTF	914	
Qy	901	QLELPVKYAVYVTVSHGVSTKYLNFTAS - ENTSRVMQHQYQVSNLQORSPLISLVFLVP	959	
Db	915	QLELPVKYAVYVTVSSHEQTKYLVNFSESEKESHVAMHRYQVNNLQORDLPVSNFPWVP	974	
Qy	960	VLRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVQRIQCDI	1019	
Db	975	VELNQEA VMDEVSHPQNSLRCSSSEKIAAPPASDFLAHQKVPVLDCSIAGCLRFRCDV	1034	
Qy	1020	PFQIQBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTE	1079	

Db	1035	PSFSVQEEELDTLKGNSLFGWVRQILQKKVSVVVAEITFTDTSVSQLPQGBAFMRAQTT	1094
Qy	1080	TKVEFEVFNPLPLIVGSSVGLLALITLALITLALYKLGFRKROYKDMWSE	1128
Db	1095	TVLEKYKVNHPPLIVGSSIGGLLLALITLALYKVGFPKROYKEMWEE	1143
RESULT 4			
S03308			
cell surface glycoprotein CD11a precursor - human			
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function			
C;Species: Homo sapiens (man)			
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 20-Aug-1999			
A;Accession: S03308; A47458; A47565; A48759; S36044			
R;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.			
J. Cell Biol. 108, 703-712, 1989			
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subu			
A;Reference number: S03308; MUID:89139587; PMID:2537322			
A;Accession: S03308			
A;Molecule type: mRNA			
A;Residues: 1-1170 <LAR>			
A;Cross-references: EMBL:Y00796; NID:g31421; PIDN:CAA68747.1; PID:g31422			
A;Note: part of this sequence was confirmed by protein sequencing			
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.			
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993			
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a)			
A;Reference number: A47458; MUID:93248261; PMID:8097887			
A;Accession: A47458			
A;Molecule type: DNA			
A;Residues: 1-20 <COR>			
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)			
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.			
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993			
A;Title: Identification of cell-specific and developmentally regulated nuclear factor			
A;Reference number: A47565; MUID:93281759; PMID:8099450			
A;Accession: A47565			
A;Molecule type: DNA			
A;Residues: 1-20 <SHE>			
A;Cross-references: GB:M95609			
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.			
J. Biol. Chem. 268, 19305-19311, 1993			
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.			
A;Reference number: A48759; MUID:93374910; PMID:8103515			
A;Accession: A48759			
A;Molecule type: DNA			
A;Residues: 1-20 <NUB>			
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406			
C;Genetics:			
A;Gene: GDB:ITGAL; CD11A			
A;Cross-references: GDB:119757; OMIM:153370			
A;Map position: 16p11.2-16p11.2			
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h			
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; t			
F;1-25/Domain: signal sequence #status predicted <SIG>			
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted			
F;154-317/Domain: von Willebrand factor type A repeat homology <VMA2>			
Query Match 26.4%; Score 1548.5; DB 2; Length 1170;			
Best Local Similarity 34.5%; Pred. No. 1.2e-98;			
Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;			
Qy	1	FNLDTENAMTFO--ENARGFGSVQVLOGSRVVGAPQEIIVAAQNGSLYQCDYSTGSCSE	58
Db	26	YNLDRVGARSFSPRAGRHFYRVLQV-GNGVTVGAPGE---GNSTGSLYQCSGTGCHCL	81
Qy	59	PIRLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTYVKGCLFLFGSLNR---	115
Db	82	PVTLR-GSNTSKYLGMTLATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFRQNLQCPM	140
Qy	116	-QOPQKPEALRCPOEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVWEOLKSKTILFSLM	174
Db	141	LQRRPGFQECIKG---NVDLVFLFDGSMISLQPDQFQKILDFMKDVMKKLSNTSYQFAAV	196

QY 763 DDLSTSFMSLDCLVGGP-----REFNVTVTRNDGDSYRTQVTFPPFLDLSYRKV 816
Db 775 ANLTLSPPARS-----GRLMSSASLAVETLNSGSDAYWRLDLDFPRGLSPRKV 827
QY 817 STLQNRQSRWRILACASSTEVSGAL-KSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCBEL-TEGSSLLTKLKNVSSPIFKAGQEVSLQWMENTLNS 882
QY 876 LGNKLKLLKANTYSEN-NMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFASNTSR 934
Db 883 WEDVELNGTVHCENENSSLOEDNSAATHIPVLVPLNQLVIMDRFO---VTFSENLS--TCHTK 986
QY 935 VMQHOYOVSNLQSRSLPISLVLVPLVRLNQVIMDRFO---VTFSENLS--TCHTK 986
Db 943 QVQHVVQV-----RIQSAVDHNMPT-LEALGVPRPHSEDLITYTWSQTDPLVTCSE 996
QY 987 E-RLPSSHDFLAELRKAPVWNCSTAVCQRIQCDIPFGIOEFNATLKGMLSPDWYIKTS 1045
Db 997 DLKRPSE--AEQCLPGV-----QFRCPVIV---RWEILLQVGTVELSKETKAS 1042
QY 1046 HNHLIVSTAILFNDSVFTLLPGCAFVRSQTEKVEPEVPNPLPLIVGSSVGGGLLL 1105
Db 1043 -STLSLCSLSVSFNSSSHFLYGSKA-SEAQVLVKVDLIHEKEMLHVYLSGGGLVLL 1100
QY 1106 ALITAAALYKLGFEKRYQKDM-SEGPGPAEP 1136
Db 1101 FLIFALYKVGFFKRNLEKMEADGGVNGSP 1132

RESULT 6
A53213
integrin alpha-E chain - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: A53213
R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. Ur
A:Reference number: A53213; MUID: 94164962; PMID: 8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: GB:L25851; NID:9457244; PID:9457245
C:Genetics:
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1149; DB 2; Length 1179;
Best Local Similarity 29.1%; Pred. No. 5.9e-71;
Matches 343; Conservative 211; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYOCYDSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPOLLACGPTVHQ 95
Db 65 GLHRCSLVQDEILCHPVEHVPIPKGRHRGVTVVRSHHGVLCI-----QVLVRP--HS 117
QY 96 TCSENTYVKGLCFLFGSNLRQOPQ-----SDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL-- 119
Db 118 LSSELT---GTCSLLGPDLRPOAQANFDLENLDPARDVDGDCYSKKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL-- 164
Db 175 RRRALKEEEDKEEEDDEEAEAGTEAIIIDGSGSIDPPDFORAKDFISNMNRNFE 234
QY 165 KSKTLFSLMQYSEFRTHFTKEFQNNPNPSLVKPIITOLLGRTHATGVRKVIKRELLN 224
Db 235 KCFECNFALVQVGGVQIEFDRDSQDVNASLARVONITQVSGVTKTASAMOHVLDISFT 294
QY 225 ITNGARKNAFKLIVITDCEKFDPLGYEDVITPEADREGVIRVYVIGVDGDAFSEKSRQEL 284

RESULT 7

A45226

integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2000

C:Accession: A45226

R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.

J. Biol. Chem. 268, 2989-2996, 1993

A:Title: Expression of native and truncated forms of the human integrin alpha 1 subu

A:Reference number: A45226; MUID:93155124; PMID:8428973

295 SSHGRRKASKVMVVLTDGGIFEDPLNLTTVINSPKMQGVVERFAIGVGEEFKSARTAREL 354
QY 285 NTIASPPRDHVPQVNNPEALKTIONOLREKIFALEGTOTGSSSSSEHEMSQEGFSAAIT 344
Db 355 NLIASDPDETHAFKVTNTYHALDGLLKLKYNIIISHEGT---VGDHALYQLAQIGFSAQL 411
QY 345 SNGP-LLSTVSGYDWMAGGVFLY-TSKEKSTFFINMTRVSDMDND-----YLGAAAIILRN 398
Db 412 DERQVLLGAVGAFDMSGGALLYLDTSRRRGRFLNQTAATAAADAEEAAQSYLYGAVAVLHKT 471
QY 399 RVQSLVLGAPRVQHIGLVAMFR-ONTGMHESNANV-KGTQIGAYFCASLCSVDVDSNGST 456
Db 472 CSLSYVAGAPQYKHG--AVFELQKEGRASEFLVLEGEOMGSGYFSELCPCVDIDMGST 529
QY 457 DLVLIGAPHVYEQTRCGQVSVCPPLPRGARWQCDVAVLGEQCPWGRFCAALTVLGDVN 516
Db 530 DFLVLAAPFVHVHGEGRVTVYRLSE-QDGSFSLARILSGHPQFTNARFGFMAAAGDLS 588
QY 517 GDKLTDAVATGAP---GEEDNR--GAVLYPHGTSGSGISPSHSQRIAGSKLSPLRYFGQ 570
Db 589 QDKLTDAVATGAPLEGFGADDGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
QY 571 SLGGQDLTMDGLVDLTGVAQGHVLLLRGQVLRVKAIMEFNPREVARNVFECDQVVGK 630
Db 648 SMAGGFDISGDGLADITVGTGQAVVRSRPPVRLKVSMAFTPSALP-----IGF 697
QY 631 KEAGEVRVCLHVQKSTRDLREQIOISVVVTDIALDSGRPHSRVAFNETKNSTRRTQVL 690
Db 698 NGVVNRLCFEI--SSVTASESLREALNFTLDVGVKQRRRLQDVSRCGLCREWS 756
QY 691 GLTQTCETLKLQLPN---CIEDPVSPVILRLNFSLVGTPLSFAFNLRPLVLAEDAQLFT 746
Db 757 SGSQLCEDL-LLMPTGELCEEDCFNSASVKVSYQL-QTPEGQTDHPQPLDRLRYTEPFAI 814
QY 747 ALPPEKNCNDNI CODDLSITFSFMSLDCLVGGPREFNVTVTRNDGDSYRTQVTF 806
Db 815 FQLPYEKACKNKLFCVAELQLA-TTVSQQELVWGLTKELTLNLTNSGSDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQNRQSRWRILACASSTEVSGALKSTSCSINHPIPPENSEVTFNI 866
Db 874 YPRNLQ-----LKRQKPPSPNIQCDPQPV---ASVLNMCRIQHPVL-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN-----MPRTNKTEFQ---LELPVKYAVYVMTSHGV 919
Db 924 VQOLEENAFNRTADITVTVTNSNRRSLANETHTLQFRHGFVAVLSKPSIMYVNTCQGL 983
QY 920 S--TKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPLVRLNQTVIWDPRQVTFSE 977
Db 984 SHHKELFLVHGEN---LFGAEYQ-----LOICVPTKRLGLQVAAVKKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIPFGIQEBFNATLKGMLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVEMHSHVSCVIA-----SDKENVTVAEIS 1073
QY 1038 FDMYIKTSHNHLIVST-----AEILFNDSVFTLLPGQAGFVRSQTEKVEPEVNPPL 1091
Db 1074 WD-----HSEELKQVTELQILGEISFNKSLYEGLENAENH--RTKITVTVFLKDEKYHSL 1125
QY 1092 PLIVGSSVGGLLLALITAAALYKLGFEKRYQKDMSE 1128
Db 1126 PIILKGVGGLLVILVILVILFKCGFFKRYQQLNLE 1162

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151 <BT>

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: Sequence extracted from NCBI backbone (NCBIP:124326)

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;142-317/Domain: von Willebrand factor type A repeat homology <VMA>

Query Match 18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 9.4e-68; Indels 195; Gaps 44;
Matches 344; Conservative 211; Mismatches 487;

QY 1 FNLDTENAMTFOENARG-FQSGVYVQL---QSGRVVVGAPQEIIVAAANQRLSLQCDYSTGS 56
DB 1 FNVVKNMSTFGPVEDMGTYTVOYENBEGKWLIGSLVGPQKRTGDVYKCPVGRGE 60
QY 57 CEP-IRLVQVVEA-----VNMSLGLSLAATTSPQALLACGPTVHOTCSENTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHLHYTTGI 119
QY 107 CFLFGSNLRQOPKFPEALRGCPQEDSDIAFLDGGSGIIPHDPRMKFEFVSTVMEQLK- 165
DB 120 CSDVSPTFQVNSIAP--VQECSTQ-LDIVIVLDGSGNSIYPWD--SVTAPLNDLLKRM 174
QY 166 -KSKTLFLSMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGR-THTATGVRKVIKRELL 223
DB 175 GPKQTQGVIOYGENWTHFNLNKYSSTEVLVAAKIVORGROTWALTGDTARKEAF 234
QY 224 NITNGARKNAFKILIVITDGEKGPLGYEDYDPEADREGVIRYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKVMVITDGEH-DNHRLKKVIOCDENIQRFSAIILGYSYNRGNLSTE 293
QY 279 KSRQELNTASKPRPHVFOVNNFEALKTIQNLQREKIFAIEGTQTGSSSSFEHMSQEG 338
DB 294 KFVEEKISASEPTEKHFFNVSDALVITVTKLGERIFALEATADQSAASFEMMSQTG 353
QY 339 FSAATISNGPLLTGVSYDAGGVFLVTSKE-----KSTF-INMTRVDSMDNDAYLGAA 392
DB 354 FSAHYSDWMLGAVAYDWNGVVMQKASQIIIPNTTFNVSTKKNPL-ASYLGYYTV 412
QY 393 AITLRNRVQSL-VLGAPRYQHIGLVAMFRONTGMESNANVKGTQIGAYFGSLCSVDVD 451
DB 413 NSATASSGDVLYIAGQPRYHNTQGVIIYRMEDGNIKIQLTSGEIGSYFGSLITTTDID 472
QY 452 SNGSTDLVLIGAPHY-----YEOTR-GGVSVCPPLPRGORARWQCDVAVLY 495
DB 473 KDSNTDILLVGA PMYMGTEKEEGKVYVVALNQTREYQMSLEPIKQTCSSRQHNSCTT 532
QY 496 GEGQOPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNARGAVYLFHGTSGSGISPSHQ 554
DB 533 ENKNEPCGARFGTAIAVNDLNDGNDIVIGAPLSDDHGGAVYIIYHG-SGKTIKRYAQ 591
QY 555 RIAGSKLSPRLQYFGSGSLGGQDLTMDGLVDLTGVAQGHVLLLRSPVLRVKAIMBPNR 614
DB 592 RIFSGDGDKTLKFGSGIHEMDLNGDGLTDTVIGLGGALPWSRDAVVKVYTMFEPN 651
QY 615 EVARNVFCNDQVVKGEAG--EVVCLHVQ-KSTRDRLEGGIQQSVVYVLDALDSGRPH 671
DB 652 KVNQKKNCH---MEKETVCINATVCFVKLAKSKEDTIEADLQ---YRVTLDSLRQI 704
QY 672 SRAVFNET-----KNSTRQTOVLGTQTCETLKLQLPNCIEDPVPVILRLNFSLVGT 725
DB 705 SRFFSGTQBRKVQNRNITVRKSEC-----TKHSFYMLDRKHDFQDSVR---ITUDFNLT-D 755
QY 726 PLISAFGNLRPLVAEDAQLFTALFPPEKNCNDNIQODLSITFSFMSLDCLVVGPRE- 784
DB 756 PENG-----PVLDDSLPNSVHEIYPAKDCGNKEKICISLSLHVAATEKDLLIVRQNDK 810
QY 785 FNVTVVRNDGDSYRTQVTFPFLDSLVRKVTSLQNRQSRWRLACESASSTEVSGAL 844
DB 811 FNVSLTVKNTKDSAYNTRTIVHYSPLNVFSGIEAIQKD-----SCSEN-----853

QY 845 KSTSCSINHPIPPENSEVTFNITFDVDSKASIGN-KLLLLKANVTSNNMPTNKTKEFOLE 903
DB 854 HMITCKGVYFLRRGEMVTFKILFQNTSYLMENVTIVLSATSDSEEPETLSDNWNVIS 913
QY 904 LPVKYAVYVMTSHGVSTKYLNFASENTSRVMQHOYQVSN-----LQORS-----L 950
DB 914 IPVKYEVGLQFYS-SASEYHISIAAANTVPEVINSTEIDIGNEINIFYLIRKSGSPMPBL 972
QY 951 PISLVF-----LVPVRLNQTVIWDPRQVTFSENLSSTCHTKE-----RLPS 991
DB 973 KLSISFPNMTSNGYVPVLYPTGLSS-----SENACRPHIFEDPFSINSCKMTT 1021
QY 992 HSDFLAEIRKAPVNCISIAVCORIQCDDIPFFGQIE-----EFNATLK 1033
DB 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFKYSFSSNLNLTIR 1078
QY 1034 GNLSPDWIKTSHNHLILVSTAEIILFNDSVFTLLFGQAGFVRSQVETKVEPEVNPPL 1093
DB 1079 GEL-----RSENASLVSSN-----QKRELAIOISKDGLPGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKRYQKMMSE 1128
DB 1115 WVILLSAPAGLLLLMLLILALWKIGFFRPLKXKMEK 1151

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C;Accession: I45914

R;Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A;Title: Identification of putative ligand binding sites within the I-domain of integrin

A;Reference number: A54402; PMID:94193847; PMID:7511592

A;Accession: I45914

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1170 <KAM>

A;Cross-references: GB:L25886; NID:9439695; PID:AA859255.1; PID:9439696

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;161-336/Domain: von Willebrand factor type A repeat homology <VMA>

Query Match 18.5%; Score 1084; DB 2; Length 1170;

Best Local Similarity 27.7%; Pred. No. 1.9e-66;

Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFO-ENARGFQSGVYVQL---QSGRVVVGAPQEIIVAAANQRLSLQYC--DYST 54
DB 19 YNVGLPKAKIFSGPSSEQFGYAVQQFIPNKGWLLVGVSPWGFPPKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQALLACGPTVHOTCSENTYVKG 107
DB 79 TTCEKLNLTSTSMNSVTEMKTNMSLGLTLTRNVGTGGFLTCGPLMAQQSGSYTYTGV 138
QY 108 FLFGSNLRQOPKFPEALRGCPQEDSDIAFLDGGSGIIPHDPRMKFVSTVMEOLK-- 165
DB 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVCDENSIYPWD--AVNLFKFXVQGLDIG 194
QY 166 KSKTLFLSMOYSEEFRIHFTFKFQNNPNRSLVKPITQLL-----GRTHATGVRKVI 221
DB 195 PKTKQMLLIQYANNPRVFNLTNFKSKD---EMIKATSTQTFQYGGDLTNTFKAIOVARDT 251
QY 222 LNIITNGARKNAFKILIVITDGEKGPLGYEDYDPEADREGVIRYVIGV-----GDAFR 276
DB 252 AYSTAAGRPGATKVMVVVTDGESH-DGSKLXAVIDQCNKNILRFGIYGLVGLYNNALD 310
QY 277 SEKSRELNTIASKPRDRHVFOVNNFEALTKIQNLQREKIFAIEGTQTGSSSSFEHMSQ 336
DB 311 TKNLKIKAIASIPTEHFHFFNVSDADLLEKAGTIGEIQFSLTEGVQO-GNFFQEMSQ 369
QY 337 EGFSAAIT--SNGPLISTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGVA 391

Db 370 VGFAEYSPQNNILMLGAVGAYDMSGTVVQKTPHGHILFQKAPQEQIQLDRNHSSYLGS 429
Qy 392 AAILLRNVQSLVLAGARYQHIGLVAMFRONTGWMESNAVY-----KGTQIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANTGTQIVLYSVN-----ENGNTVTIQSORGQDQIGSYFQSVL 484
Qy 446 CSVDVDSNGSDTLVLIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDVLYGEQGPWG 503
Db 485 CAVDNKTITDVLVAGAPMYMNDLKKBEGRVYLTITKG-ILNWH--QFLEGNGLENA 541
Qy 504 RFGAALTVLGVNGDKULTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSGRIAGS--KL 561
Db 542 RFGSAIAALSIDINMGDFNDVTVGSPLENQNSGAVIYNGHEGM-IRLRYSQKILGSDRAF 600
Qy 562 SPRLOYGQSISGGDLTMDGLVDLTGACQGHVLLRSOPVLRKATMEFNPREVARNVF 621
Db 601 SSHLOYGERSLDGVDGLNGDSITDVSGAFQGVQVQLWSQSIADVSDASTFPKXI--TL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVTVYDLALD----SGRPHSRAVEN 677
Db 659 NKNAEI-----KLKLCF-----SAKFRPTNQNNQVAIVYNTIDEDQFSSRVSRLGFK 707
Qy 678 ETKNSTRQTVGLGTQCE--TLKLQLPNCIEDPVPVILRLNPSL--VGTPLSAPGNL 733
Db 708 ENNERCLOKTMVSAQORCSEYIIHIOEPS---DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVAEDAQRALTALFPEKNCNGDNTCODDLSTF---SFMSLDCLVGGPREFNVTV 789
Db 757 NPALAEVSETVKVFSIPPHKDDGGDGVCSIDLVLNVQQLPATQQOPFIVSNQNLRTFSV 816
Qy 790 TVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORSQSRWLACESASST-EVSGALKSTS 848
Db 817 QLNKKSAYNTEIVDFSENLF-----ASMSMPVDGTETVCQIASQSKSVT 864
Qy 849 CSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEFQLELPVKY 908
Db 865 CNVGPALKSQQVTFITNFDNLQ-ILNQOASISFRALSQBENMADNSVNLKLSLLY 923
Qy 909 AVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQOR-----SLPISLFLV 958
Db 934 DAETHIT-RSTNINPVEVSLDGNVSSVY-HSFE--DIGPKFIPSKVTTGVSVPVSM- 976
Qy 959 PVRLNQTIVDRPQVTFSEN---LSSTCHTKE-----RLPSSHDFLAE- 998
Db 977 -----SVIHIPTKDKNPLMYLTGVHTQAGDISCEAEINPLKIGQTSVSSVFSKEN 1030
Qy 999 LRKAPVNCSTAVCORIQCDIPFGIQBEFNATLKNLSFDWYIKTSHNHLILVSTAEI- 1057
Db 1031 FRHIKELNCRTASCNSINCMWLRDLQVKGFEFLNVSTRIWNGTFAASTFQTVLQTAABID 1090
Qy 1058 LPNDSVFTL-----LPCQAGFVRQSOTETKVEPEE-VPNPLPLVIGSSVGLLALLITA 1110
Db 1091 TYNPOIVYIENTVTIP-----LTIMKPEKVEVPTGVIGSVIAGILLALLVA 1140
Qy 1111 ALYKLGFFKRYOKOM 1125
Db 1141 ILMKLGFFKRYOKM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A;Cross-references: EMBL:229987; NID:g473098; PIDN:CAA82877.1; PID:g473099
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
P:169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1072; DB 2; Length 1178;

Best Local Similarity 28.1%; Pred. No. 1.3e-65;

Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSVVVVGAPQEI VAAHQRLSYOC--DYST 54
Db 27 YNVGLPKAIFSGPSSEQFGYSVQQLTNQGNWLLVGSFWSGPFENRMDGVYKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSAAATTSPPQLACQPTVHQTCSNTYVKGIC 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCPLWAHQCGNYYATGIC 146
Qy 108 FLFGSNLRQPOQ---KFPALRCPCQEDSDIAFLIDGSGSIIIPDFFRMKEFVSTWEOQL 164
Db 147 ----SDVSPDFQFLTFSFAVQACPSL-VDVVVVCDENSIYP--WEAVKNFLVKFVTGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTFKBFQNNPNRSLVKPITQLLG-RTHTATGVKRVIRE 221
Db 200 DIGPKKTOVALIOYANEPRILFNLDNFETKEDMVQATSETROHGGDLTNTFRAIEFARDY 259
Qy 222 LNMITNGARKNAFKILIVITDGEKFGDPLGYEDVIPLEADREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSGRPRCATKVMVWVTDGESH-DGSKLKTVIQOCNDEILRFGIAVLGVLLNRNALD 318
Qy 277 SEKSRQELNTIASKPRDRHVQVNNFEALQIONREKIFAIEGTQTGSSSFEHMSQ 336
Db 319 TKNLKIKETKAIASITPTERYFFNVADAEALLEKAGTIGEQIFSTEGTVQG--GDNFQMBMAQ 377
Qy 337 EGFSA--AITSNGPLSLTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSMN-DAYLGVA 391
Db 378 VGFSAVAPQNDILMLCAVGAPFWSGTLVQETSHKPVIFPKQADFQVLDQRNHSSFLGYS 437
Qy 392 AAILLRNVQSLVLAGARYQHIGLVAMFRONTGWMESNAVY-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHVFAGAPRANTGTQIVLYSVNK---QGNVTVIQSHRGDQIGSYFVGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDVLYGEQGPWGRF 505
Db 495 VDVKDQITITDVLVAGAPTYNNDLKBEGKLYLTITKILNQHQ---FLEGPEGTGNARF 551
Qy 506 GAALTVLGVNGDKULTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSGRIAGSKLSPR- 564
Db 552 GSAIAALSIDINMGDFNDVTVGSPLENQNSGAVIYNGHQST-IRTKYSQKILGSGAFRR 610
Qy 565 -LQYFGQSLSGGQDLTMDGLVDLTGACQGHVLLRSOPVLRKATMEFNPREVARNVFEC 623
Db 611 HLQFFGRLSDYGLNGDSITDVSIGALGQVQLWSQSIADVAIEALPTP----- 660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVTVYDLALD----SGRPHSRAVENET 675
Db 661 -DKITLLNKDAKITLCLCFRAEFPAGQNNQV--AILFNMTLDADGHSSRVTSRGVREN 717
Qy 680 KNSTRROTQVGLTQTCT--LKLQLPNCIEDPVPVILRLNPSLVLGTPLSAFGLNLPVL 737
Db 718 SERFLQNMVNVQKCEHHISIQKPS---DVNPLDLRVDLSLENPGTS-----PAL 768
Qy 738 AEDAQLFTALFPPEKNCNDNICQDDLSI-----TFSPMSLDCLVGGPREFNVTV 788
Db 769 EAYSETVKVFSIPFYKEGSGDGCISDLIDVQQLPAIQTSF-----IVSNQNLRTFS 823
Qy 789 VTVNDEGDSYRTQVTFPPFLDLSYRKVSTLQORSQSRWLACESASST-EVSGALKST 847
Db 824 VILKNRGEYNTVLAEPFSENLF-----ASFSPVDGTEVTCEVSSQKSV 871
Qy 848 SCSTNHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRNTKTE--FQLELP 905
Db 872 TCDVGYPALKSEQQVTFITNFDNLQ-ILNQOAAINFQAFSESO--ETNKADNSVSLTIP 928
Qy 906 VKYAVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQOR-----SLPISLV 955


```

Db 929 LLYDAELHLT-RSTNINFEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSNMA 984
QY 956 FLV-----PVRLOTWIDRPOVTF--SENLS 980
Db 985 LVTIHIPOYTKENKPLLYLTGTQDQAGDISCTAEINPLKLPHTA-----PSVSPKNEPR 1040
QY 981 STCHTKERLPSSHDFLAELRKAPVNCVSIACQRIQCDIPFGIOBEFNATLKGNSLFDW 1040
Db 1041 ---HTKE-----LDCRTTSCNITCMLKDLHMAEYFINTVTRVWNR 1080
QY 1041 YIKTSHNHLIIVSTAEILNDSVFTLLPQGFAPVRSQTEKTPPEVEPNPLPLVGVSSVG 1100
Db 1081 FAASTFTQVQLTAALAEIDHNPOLFVIEENAVTIPLIMKPTKEAEVPT--GVIIIGSIIA 1138
QY 1101 GILLALLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLALLAMTAGLWKLGFQYKDM 1163

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 21-Jul-2000
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A:Reference number: A33998; PMID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <AK>
A:Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A>Note: the authors translated the codon GAT for residue 802 as Gin, GTC for residue 803
Biochem. J. 279, 419-425, 1991
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC*, GPIIa an
A:Reference number: A56793; PMID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A:Reference number: A53117; PMID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A>Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
A:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,343,432,460,475,699,1037,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 2,1e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFO-ENARGFGQSVQL-----QGRVVVVGAPQETVAANQKSLVQC--DYST 54
Db 30 YNVGLPEAKIFGSPSQEQGYAQVQFINPKGNWLLVGSPPFNRMGDVTKCPVDLST 89

```

```

QY 55 GSCPIRLQ-----VPEVAVNMSLGLSLAATTPPOLLACGPVTHQTCSENTYVKGLC 107
Db 90 ATCEKLNLTSTSPNVTETKNTNMSGLILTRNMGTCGFLTCGPPLWACQCGNOYVTTGVC 149
QY 108 FLFGSNLBOQPKPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKKFPVSTVMEQLK-- 165
Db 150 SDISPDF--QLSASFPATQPCPSL-IDVVVVVCDSESNISYPWD--AVKNFLEKFKVQGLDIG 205
QY 166 KSKTLFSLMQYSEBFRIFHTFEKFNPNPRSLVKPITQLLG-RTHATATGVKVRRELLN 224
Db 206 PTKQVGLIQAANNPRVFNLTYYTKEEMIVATSQTSQYGGDLTNTFGAIOYARKAYS 265
QY 225 ITNGARKNAFKILIVITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAPRSEK 279
Db 266 AASGRRSATKVMVWVTDGESH-DGSMKAVIDQCNHNIILRFGIAGVLGNALNADTKN 324
QY 280 SRQELNTTASKPPRDHVFQVANNFEALKTIONLBEKIPAIETGTQTSSESSSEHEMSQSGF 339
Db 325 LIKEIKATIASPTERYFFNFVSDAALKEKAGTLGEQIFSIETVQG-GDNFQEMESQYGF 383
QY 340 SAAITSNGP--LLSTVGSYDWAGVFLYTSKEKSTFINMT--RVDSDMN-DAYLCYAAAI 394
Db 384 SADYSSQNDILMLGAVGAFGSGTIVQKTSHGHLIPKQAFDQILQDRNHSYLGYSVAA 443
QY 395 ILNRVQSLVILGAPRYQHIGLVAMPQNTGMWESNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFVAGAPRANVTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498
QY 449 DVDSNGSTDVLIGAPHYYEOTR--GGQVSVCPILPRGORARWQCDAVLYGSGQGWGRFG 506
Db 499 DVDKDTITDVLVVGAPMYMSDLKKEBGRVYLFTIKKILGQHQ-----FLEGPEGIENTRFG 555
QY 507 AALTVLGVGVNGKLTVDVAIGAPGEDNRGAVYLPHTGTSGSGISPSHSORIGS--KLSPR 564
Db 556 SAIAALSINNDGNDVIVGPLENONGAVYIINGHGT-IRTKYSKILGSDGAPRSH 614
QY 565 LQYFQSLSGGQDLTMDGLDVLTVGAQGHVLLLSQPVLKVAIMEFNPREVARNVFECN 624
Db 615 LQYFGSLDGYGLDNGDSITDVSIGAFGVVQLMSQSIADVAIEASFTEKI--TLVNKN 672
QY 625 DVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVYTYDLAD-----SORPHSAVFNETHK 680
Db 673 AQII-----LKLCF-----SAKFRPTKQNNQVAIVYNTILDADGFSRVRTSRGLFKENN 721
QY 681 NSTROTQVGLGTQTC--ETLKLQLPNCIEDPVPSPVLRLNFSLVGTPLPSAFGNLRPVLA 738
Db 722 ERCLQNMVNVQAQSCPBHIIYIQPS---DVVNSLDLURVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALFPFKEKNCNDNICODDLGITF-----SFMSLDCLVYVGGPREFNVTVTRND 794
Db 773 AYSETAKVSPFPHKDCGEDGLCISDLVDVRQIPAAQEPPIVSNQNKRLTFSVTLKXK 832
QY 795 GEDSYRTQVTFPFLDLSYKRVKSTLQONORSORSHLACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDVFSENLFF-----ASFSLPVDGTEVTQVAAASQKSVACDVGY 880
QY 854 PIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFOLELPVKYAVMV 913
Db 881 PALKREQVQVTFINPFDNLQ-NLQNAQSLSPALSESQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSHGVSFKYLNFTASENTSRVMQHOYQVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLTRSTNINFVEISSDGNVPSIVHSFEDVGPKFIFSLKVTTCGSPVSMATVILHPQ 995
QY 959 -----PVRLNQTVIWRPQVTF--SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGT-----SSVSFKSENPR---HTKE- 1047
QY 998 LPSHSDFLAELRKAPVNCVSIACQRIQCDIPFGIOBEFNATLKGNSLFDWYIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCWLKDVHMKGEYFVNVVTRINWGTFFASSTFOT 1091

```


Db 101 RKNPQTCBOLQSGSPGCPGKTCLEERDQWLVTLRSQPCGNGSIVTCGHRWKNIPY 160
QY 96 TCSNTYVYKGLCFGLFSGNLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGIIPHDPFRMK 155
Db 161 MKSDNKLPTGICVMPDRLTSLK-----185
QY 156 FVSTVMEQLKSKTFLPSLMQYSEEPRIHTFKEFQNNPNRSLVKPITOLLGRTHTATGV 215
Db 186 -----RMAPCYKDYT-----195
QY 216 RKVIRELLNITNGARKNAKILIVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDAF 275
Db 196 -----RKFG-----200
QY 276 RSEKSRQELNTIASKPPRDHVFQVNNFEALKTIONLREKIFAIEGTQTGSSSSPEHMS 335
Db 201 -----NFAS-----C 205
QY 336 QGFSAAITNSNPLSTVSGSYDWAGGVFLY--TSKEKSTFINMTRVDSMDNDAYLGYA-- 391
Db 206 QAGISSFTQDLIVMGAPGSSYMTGTVFVYNTTNOYKAFVD--RQNVKFGSYLGSYVG 263
QY 392 AAILLRNVQSLVGLAPRYQHICGLVAMFRONTGMWESNANV-----KGTOIGAYFGASLCS 447
Db 264 AGHFRSPHTEVVGAPQHEQIGKAYIF-----SIDENELNIVYEMKKGKLSYFGASVCA 319
QY 448 VDVDSNGTDLVLIGAPHYVEQTRGGQSVCPPLRGORA--RWQCDALVYGEQGPWGRFG 506
Db 320 VDLNADGFSDL-LVGAPMSTIREGRVFVY-INSNGAVVMEMERVLVGSBKYA-ARG 376
QY 507 AALTVLGVNGDKLTDVAIGAPEEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLRQ 566
Db 377 ESIANGLDINDGFEFEDIALGAPQEDDLRGAVIYNGRV-DGISSTYSQRIEQQISKSLR 435
QY 567 YGQSLSGQDLTMDGLVLTGVA--QGHVLLRSQPLRVKAIEMFNPREVARNVFECN 624
Db 436 MFGQSISGQIDADNNGYGVDAVGAPOSASVALLRTPVIVVEASLS-HPESVNRRTKFDCT 494
QY 625 QDVVKGKEAGEVRVCLHVOKSTRDLREGOIQSVVTDIALDSGR---PHSRAVF--NBT 679
Db 495 -----ENGLPSVCMHLTCLFSYKGEVPGYIVLVYNSLDVHRKAESPFRYFSNGT 547
QY 680 KNSTRQTVGLGTQTCETLKLQPLNCIEDPVSPIVLRNFSI-----VGTPLSAFNGLR 734
Db 548 SDVITGSIRVSSGKCRTHQAPMRKDVRLDILTPIHVEATYHLGHVITKRNTEEPPLQ 607
QY 735 PULAEADAQR-LFTALFPPEKNCNDNICODDLSITFS-----FMSLDCLVYGGPREFN 787
Db 608 PILQOKKEDVIRKMINFARFCAVEN-CSADLQVSAKVGFLPKPYENKTYLAVGSMKTI 666
QY 788 TVTVRNDGDSYRTOVTFEFPDLDSYRKVSTLQNRQSRWRLACESASSTVSGALKST 847
Db 667 NVSLFNAGDVAETTLNVQLPTGLYFIKILDUDEK-----QINCE---VTSSGLVK-L 716
QY 848 SCGINHPFPENSEVTNFTFDVDSKASLGNKLLKANVTSEN--NMPRTNKTEFOLELP 905
Db 717 ACSLGYIYVDRLSRIDISLLDVSSLSRAHEDLSISVHASCENEGELDQVRNVRTLTIP 776
QY 906 VKYAVVMVTSRGV--STKYLNFPTASENTSRVNHQO-----YOVSNLQORSPL-ISLVL 957
Db 777 LRYEV--MLTVHGLVNPSTFVYSGSSENEPETCMAEKLNLTHVINTGISMAPNVSVKIM 834
QY 958 VPVRLNQTVMWRP--QVTFSENLSSTCHKE-----RLPSSHSDLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLDVQTTTCQCHFGYGRECTFAQQKGIAGTLTDIVKFLSKTD 991
QY 1001 KAPVNVCSIA--VCQRIQCIDPFGIQEEFNATLKNLSFDWYIKTSHNHLITVSTAEIL 1058
Db 892 KR-LLYCMKADQHCDFLCN---FGKWESG-----KEASVHIQLEGPSIL 933
QY 1059 FNDSVFTLLPGQAFVRSQETKVEPFEVFPN-----1090
Db 934 EMDETSSL-----KFEIKATAPPEPHPKVIELNKDENVAVHVFLEGLHHQRKHF 983

QY 1091 -LPLIVSGSVGGLLLALITAAALYKLGFFKQYKDMMS 1128
Db 984 TIIITISLLGLVLLISCVMMKAGFFKQYKSILOE 1022

RESULT 13
S06046

integrin alpha-4 chain precursor - human
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 24-Sep-1999
C:Accession: S06046; A39355; D28018
R:Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other integrins
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046
A:Molecule type: mRNA
A:Residues: 1-1038 <TA3>
A:Cross-references: GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1; PID:g33946
R:Roehn, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: GB:M62841
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A:Title: The very late antigen family of heterodimers is part of a superfamily of mole
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018
A:Molecule type: protein
A:Residues: 40-50, 'E', 52-53 <TA2>
C:Genetics:
A:Gene: GDB:ITGA4; CD49D
A:Cross-references: GDB:128032; OMIM:192975
A:Map position: 2q31-2q32
C:Superfamily: integrin alpha-4 chain
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 633; DB 2; Length 1038;
Best Local Similarity 25.2%; Pred. No. 2.8e-35;
Matches 248; Conservative 155; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAFRSEKSRQELNTIASKPPRD-----HVFQVNNFEALKT-----IQN 310
Db 121 GKTCLLEERDQWLVTLRSQPCGNGSIVTCGHRWKNIFVKNENKLPCTGCGYGVPPDLRT 180
QY 311 QUREKI-----FAIEGTQTGSSSSPEHMSQEGFSAAITNSGPLLTSTVGSY 356
Db 181 ELSKRIAPCYQDYVKKFGENFA-----SCQAGISSFTYTKDLIVMGAPGSS 225
QY 357 DWAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGYA--AAILLRNVQSLVGLAPRYQ 411
Db 226 YMTGSLFVYNTTNYKAPLDKQNVK---GSYLGYSVGAGHFHSQHTTEVVGAPQHE 282
QY 412 HIGLVAMFRONTGMWESNANV---KGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 467
Db 283 QIGKAYIP-----SIDEKELNIIHEMKKGLSGYFASVCAVDLNADGFSDL-LVGAPMOS 337
QY 468 EOTRGQVSVCPPLRGQRRMOC--DAVLYGEQGPWGRFGAALTVLGVNGDKLTDVAIG 526
Db 338 TIREGRVFVY-INSNGAVVMAMETNLVGSBKYA-ARFGESIVNLGDIINDGFEFEDVIG 395
QY 527 APGEEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLRQYFGQSLSGQDLTMDGLV 586
Db 396 APQEDDLOGAIYIYNGRA-DGISSTFSQRIEGLSKLSLMSFMFGQISGQIDADNNGYVDV 454

QY 587 TVGA--QGHVLLRSQPLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQK 644
 Db 455 AVGAFRSDSAVLLKTRFVPIVDASLS--HPESVNRKTFDC-----VENGWSPVCIDLT 506
 QY 645 STRDLREGQIQSVVTVYDLALDSGR-----PHGRAVNETKNSTRRTQVTLGLTQTCETL 699
 Db 507 CFSYKGEVPGYIVLFYNNMSLDVNRKAESPFRFSSNGTSDVITGSIQVSSREANCRTH 566
 QY 700 KLQLPNCIEPVPSPVIRLNFSLVGTPLS-----AFGNLPPVLAEDAQR--LFTALPFPFK 733
 Db 567 QAFMRKDVDRDILTPIQIEAAYHLGPHVVISKRSTEEFPPLQPILOQKEKIMKKTINFAR 626
 QY 754 NCGNDNICQDLSIT--FSFM-----SLDCLVGGSPREFVTVTVRNDEGDSYRTQVTF 807
 Db 627 FCAHEN--CSADLQSAKIGFLKPHENTYLAAGSMKTLMLNLSFNAGDDAYETTLHVKL 685
 QY 808 PLDSYRKVSTLQNRQSRWRSLACESASSTEVSGALKSTSCSINHPIPFENSEVFNIT 867
 Db 686 PVGLYFIKILELEK-----QINCE--VTDNSGVVQ--LDCSIGIYVVDHLSRIDISPL 735
 QY 868 FVDVSKASLGNKLLKANVTSEN--NMPTNKTEFQLELPVKYAVYVWTVSHGVSTKYL 925
 Db 736 LDVSSLGRABEDLSITVHATCENEEDNDLKHRSVTVAIPLKYEVKLTVHGFVNPTSFVY 795
 QY 926 PTASENTSRV-----MQHQVQVSNLQORSLP--ISLVFLVPVRLNQTVIMDRPOVTFSEN 979
 Db 796 GSDNEPETCWEMKMLTHVINTGSMAPNVSEIWPNSFS-----PQDCLKFNI 848
 QY 980 -----SSTCHTKERLPSSHDSFLAELRKAPVWNCIAVCQRIQCDIPFGIQEBFNATLK 1033
 Db 849 LDVQTTTGECHFENYQ-----RVCALEQ-----QKSAMQTLK 880
 QY 1034 GNLSFDWYKTSNHLIIVSTAE-----ILFN-----DSVFTLLPGGAFVRQTE 1079
 Db 881 GIVRF--LSKTRDRLLYCYKADPHCLNFCNFGCMESGKEASVHIQLEGRPS--ILEWDE 936
 QY 1080 TKVEPFEV-----PNP-----LPLVGSVGGLL 1103
 Db 937 TSALKFEIRATGFPPEPNRVLNKNDENAHVHLEGLHQRPKRYFTIVISSLLGLI 996
 QY 1104 LLALITAAALYKLGFFKQYKDMSE 1128
 Db 997 VLLISYVMWKAGFFKQYKSILQE 1021
 RESULT 14
 I58409
 integrin alpha-9 chain precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
 C:Accession: I58409; A49459
 R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
 Oncogene 9, 611-619, 1994
 A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
 A:Reference number: I58409; MUID:94119603; PMID:8290272
 A:Accession: I58409
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1035 <RES>
 A:Cross-references: GB:D25303; NID:G464180; PIDN:BAA04984.1; PID:G533327
 R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
 J. Cell Biol. 123, 1289-1297, 1993
 A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
 A:Reference number: A49459; MUID:94064789; PMID:8245132
 A:Accession: A49459
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 30-1035 <PAL>
 A:Cross-references: GB:L24158
 C:Superfamily: integrin alpha-4 chain
 C:Keywords: glycoprotein; metal binding; transmembrane protein
 F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
 Best Local Similarity 26.5%; Pred. No. 4 5e-35;
 Matches -237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;
 QY 331 EHEMSQEGSAITSNTSLLSTVGSYDAGGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389
 Db 190 EHGSCQAGIAGFTTEELVVMGAPGSPFYWAGTIKVLNLTQ--NTYKLNDEVINRRRTYILG 248
 QY 390 YA--AAIILNRVQSILVLCAPRYQHIGLVAMER--QNTGMWESNANVVKTOIGAYGASL 445
 Db 249 YAVTAGHFHPSTIDVVGAPQDKGIGKVIYIFRADRRSGTLIKIPOAGKMGWSYGSLS 308
 QY 446 CSVDVDSNGSTOLVILGAPHYVEQTRGGQVSVCPPLRGQARWQCDAVLVYGGQGPWGRF 505
 Db 309 CAVDLNGDGLSL-LVGAEMFSEIRDEGQVTV--INRNGALEE--QLALTGD--GAYNAHP 364
 QY 506 GAALTVLGVNDGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPL 565
 Db 365 GESIASLDDLDNDGDFDVAIGAPKEDDFAGAVYIYHGDAG--GIVPQYSMKLSQKINPVL 423
 QY 566 QYFGQSLSGGDLTMDGLVDLTGVA--QGHVLLRSQPLRVKAIMENPREVARNVPEC 623
 Db 424 RMFGQSIGGIDMDGNGYPDVTVGAPMSDSVLLLRAPVITVD--VSIFLPGSINITAPQC 482
 QY 624 NDQVVKGEAGEVRV--CLHVQ-----KSTRDLREGQIQSVVTVYDLALDSGRPHSRV 675
 Db 483 HD-----GQPVNCLNVTTCFSFHGRKVPBEI---GLNYVLMADVAKKEGQMPRVY 531
 QY 676 F---NETKNSTRQTVGLTQTCETLKLQLPNCIEDVPSPVIRLNFSL-----VGTPLS 728
 Db 532 FVLLGETMGQVTEKLTQTYMEETCRHYVAHVRRVQDVISPIVFEAAYSLSSEHVTGEER 591
 QY 729 AFGNLRPVL-----AEDAQLFTALPFPKNCNDNIQDLSITFSM--SLD-----CL 777
 Db 592 ELPLPTVLRWKKGQKIAQKNQTV---FERNCRSED--CAADLQGLKLLSSMDEKXTLYL 647
 QY 778 VVGGRPEFNVTVVRNDGEDSYRTQVTFPFPPLDSYRKVSTLQNRQSRWRSLACESAS 837
 Db 648 ALGAVKNIISLNIISILGDDAYDANVSFNVSRLEFF-----INMQKEEMGISCLEL 701
 QY 838 TEVSGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
 Db 702 DFL-----KCSVGFPPFMRKSKYEFVIFDTSHLS--GEEVLSFIVTAQSG--NTER 750
 QY 898 TE-----FOLELPVKYAVYVWTVS-----HGVSTKYLNTASENTS---RVMQHOY 940
 Db 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFVYGESVDAANFIQDLDLECHFPQINIL 810
 QY 941 QVSNLQORSLPISLVFL--VPVRLN-----QTVIMDRPOVTFSENLSSTCHTKER 988
 Db 811 QVNTGPTSLPGSSVSIISFPNRLSSGGAEMFHVQEMVVGQEKNCFSQKQPTCIIPQEQ 870
 QY 989 LPSHSDFLAELRKA-----PVVNCIAVCQRIQCDIPFGIQEBFNATLKNLSFD 1039
 Db 871 ENIFHTIFAFFTKSGRKVLDCXPKGISCLTAHCN-----FSALAKEESRTI-----D 917
 QY 1040 WYIKTSNHLIIVSTAEILFNDSVFTLLPGGAFVRQTEKVEP----- 1084
 Db 918 IY-----MLLNT--EILKKDSSVIQ-----FMSRAKVKVDPALRVVVEIAHGNPEV 962
 QY 1085 ---FEVFN--PLPLVIG-----SSVGGLLLALITAAALYKLGFFKQYKDM 1126
 Db 963 TVVFEALHNLNPRGVVGVGIIAISLLVIGILFILLVLLWKGFFRRRYKEII 1015
 RESULT 15
 T31437
 integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
 C:Species: Lytechinus variegatus (variegated urchin)
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T31437
 R:Hertzel, P.L.; McClay, D.R.
 submitted to the EMBL Data Library, May 1998

A:Description: Alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035

A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HER>

A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 5.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

```
QY 336 QEGFSAATSNGP--LLSTVSYVDWAGGVFLYTSKEKSTFINNTRVDS-----DMNDAYLG 389
DB 181 QAGFGIIFSDNSALVNGAPGSYYLQGIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
QY 390 YAAAI--ILNRVQSLVLGAPRYOHI-GLVAMERQNTGMWESNANVKGTOIGAYFGASLC 446
DB 240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLNQFN---QWGTQIVAYFGYSVT 296
QY 447 SVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCELPGRQARWQCDVLYGEQ-----498
DB 297 VVDI--NNDTVDDLVLGAPMTDDGPAIQ-----RWEAGAVVYVQLNPDPVGPGA 343
QY 499 -----QOPNGRFGAALTVLGDVNGDKLTVAGAPCEENRGAVYLFHGTSGS 546
DB 344 SNRLSLSTLIGQIIRSRFGLSASIGDSNQDGFNDVAIGAPYEGDDAGAVIYHG--SAN 402
QY 547 GISPSHSQRIAGSKLS-PRLOVFGQSLSGQDITMTDGLVDLTVGAO--GHVLLRSQPV 603
DB 403 GLKSTPAQVLTPTLGHSGITTFGSLQGGQDMKNKYPDLLVGAESANTAVLIRPVV 462
QY 604 RVKAIMFNPREVARNVFECNDQVKGKEAGEVRVCLHVQKSTRDLREGIQSV-----658
DB 463 SLDATLNTPE-----IGINLENKTYE-LADGTMVTSFIAMT 497
QY 659 -----VYYDLALDSG-RPHSRVFNETHNSTRROTQVIGL-TQTCETLKLQ 702
DB 498 CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMEFVNDMSIITKTRRLAVSTQFCDFLRAY 557
QY 703 LPNCIEDPVSPVLRNFSLVGTPLSAFGN-----LRPVLAEQAQLFTALFPFEK 753
DB 558 VGNISIEDKLTPIKVTLOYDL-----NDESRLQPHETLPIIDMATMSTQTKQVSIQN 609
QY 754 NCGNDNICQDDLSTIFSMSLDCLVGGPREFNVTVTRNDGDSYRTQVTFPPFLDLSY 813
DB 610 NCYN-NICIPDLDTVT-PNLNPIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLGLQF 667
QY 814 RKVSTIQNQRQSRWLACESASTEVSGALKSTSCSINHPIFPEH-----SEVTNIT 867
DB 668 VRL-----ERKANMDFSVTCSQSD-----LRIITCDTGNPMVGNILFGLTSTFOVS 717
QY 868 FDVDS-----KASLGNKLLKANVTSNNMPRINKTEFQLELPVKYAVYVMTSHGVSTK 922
DB 718 GDKDSIEFYFKAESSENS--EDPNTLNNELNMTVPVTVDCITKLLGASYPEIVMYSTQED 775
QY 923 YL-----NFTASENTSVMQHOYQV-----SNLQORSLPIS-----LVFLVPV 960
DB 776 YVVPFPKAKNASADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKQKNEGEYLFYLLGI 834
QY 961 RLNQTVIWRDPQ-----VTFSENLSSTCHTKERLPKSHDFLAELRKAPV 1005
DB 835 MTEEGVTCQLTQOKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892
QY 1006 NCSIAVCQRIQCDDIPFGIOEEFNAT-----LKNLSFDWYIKTSNHLIIIVSTAEILF 1059
DB 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRF---W-----ERTF 928
QY 1060 NDSVFTLLPGGAFVRSQTETKVE--PFEVPNP-----LP-----1092
```

Search completed: November 25, 2003, 14:22:09

Job time : 19.9579 secs

DB 929 QKAVSELTPVVOQTATSSASAAVKTIPYNIPLPDRFSDSTKASTLVTTEELVPPVPTPIAM 988

QY 1093 -LIVGSSVGGLLLLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1132

DB 989 WIIIVSVLGGIILLIILGLWKCGFFERKKPGEEKYAPVASADKGGP 1038

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFOENARGFCQ.....FKRQYKDMMEGGPGCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852.5	99.6	1152	1	ITAM_HUMAN
2	4476	76.2	1153	1	ITAM_MOUSE
3	3473	59.1	1163	1	ITAX_MOUSE
4	3417	58.2	1162	1	ITAD_HUMAN
5	1552.5	26.4	1170	1	ITAL_HUMAN
6	1533.5	26.1	1183	1	ITAL_MOUSE
7	1157.5	19.7	1167	1	ITAE_MOUSE
8	1149	19.6	1179	1	ITAI_HUMAN
9	1102.5	18.8	1151	1	ITAH_HUMAN
10	1093.5	18.6	1189	1	ITAH_MOUSE
11	1084	18.5	1170	1	ITAZ_BOVIN
12	1072	18.2	1178	1	ITAZ_MOUSE
13	1069	18.2	1181	1	ITAZ_HUMAN
14	1065.5	18.1	1167	1	ITAG_HUMAN
15	1060	18.0	1180	1	ITAI_RAT
16	663	11.3	1039	1	ITAA_MOUSE
17	633	10.8	1038	1	ITAA_HUMAN
18	630	10.7	1035	1	ITAG_HUMAN
19	593.5	10.1	1032	1	ITAA_XENLA
20	571.5	9.7	1066	1	ITAJ_CRISP
21	567.5	9.7	1053	1	ITAJ_MOUSE
22	555.5	9.5	1053	1	ITAJ_MOUSE
23	546.5	9.3	1050	1	ITAS_XENLA
24	543.5	9.3	1034	1	ITAV_CHICK
25	538.5	9.2	1130	1	ITAE_HUMAN
26	535	9.1	1044	1	ITAV_MOUSE
27	532	9.1	1049	1	ITAS_HUMAN
28	531.5	9.0	1066	1	ITAJ_HUMAN
29	530.5	9.0	1072	1	ITAG_CHICK
30	526	9.0	1048	1	ITAV_HUMAN
31	517	8.8	1044	1	ITAB_CHICK
32	512.5	8.7	1091	1	ITAE_MOUSE
33	498	8.5	1179	1	ITAT_MOUSE

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.U.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

ALIGNMENTS

34	494	8.4	1396	1	ITA2_DROME
35	491.5	8.4	1146	1	ITA1_DROME
36	489.5	8.3	1039	1	ITAB_HUMAN
37	489	8.3	1025	1	ITAB_MOUSE
38	489	8.3	1033	1	ITAB_MOUSE
39	486	8.3	126	1	ITAM_CAVPO
40	472	8.0	1181	1	ITAM_HUMAN
41	469.5	8.0	1106	1	ITAT_RAT
42	466	7.9	1226	1	PAT2_CAEEL
43	445.5	7.6	1139	1	INA1_CAEEL
44	423	7.2	1115	1	ITA3_DROME
45	383	6.5	1000	1	ITA5_DROME

P12080	drosophila
Q24247	drosophila
P08514	homo sapien
P53708	homo sapien
O9QUM0	mus musculus
P15578	cavia porce
Q13683	homo sapien
Q63258	rattus norv
P34446	caenorhabdi
Q03600	caenorhabdi
Q43386	drosophila
Q9W1M8	drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=920731318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]

RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=92144986; PubMed=1346576;
RX Pahl H.L., Rosmarin A.G., Tenen D.G.;
RA "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]

RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=353202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
conservation across species and homology to platelet IIB/IIIA.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]

RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171459; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?";
RL Structure 3:1333-1340(1995).
RN [11]

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment.";
RL Structure 6:923-935(1998).
RN [12]

RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
RN [13]

CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFWA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFWA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PRO; NOTE=CD guide Cdl1b entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdl1b.htm".
RN [14]

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
RN [15]

CC EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; J02227; AAB24821.1; -
DR EMBL; S22152; AAB24821.1; JOINED.
DR EMBL; S22153; AAB24821.1; JOINED.
DR EMBL; S22154; AAB24821.1; JOINED.
DR EMBL; S22155; AAB24821.1; JOINED.
DR EMBL; S22157; AAB24821.1; JOINED.
DR EMBL; S22159; AAB24821.1; JOINED.
DR EMBL; S22161; AAB24821.1; JOINED.
DR EMBL; S22164; AAB24821.1; JOINED.
DR EMBL; S22165; AAB24821.1; JOINED.
DR EMBL; S22167; AAB24821.1; JOINED.
DR EMBL; S22169; AAB24821.1; JOINED.
DR EMBL; S22170; AAB24821.1; JOINED.
DR EMBL; S22173; AAB24821.1; JOINED.
DR EMBL; S22174; AAB24821.1; JOINED.
DR EMBL; S22180; AAB24821.1; JOINED.
DR EMBL; S22181; AAB24821.1; JOINED.
DR EMBL; S22184; AAB24821.1; JOINED.
DR EMBL; S22189; AAB24821.1; JOINED.
DR EMBL; S22191; AAB24821.1; JOINED.
DR EMBL; S22192; AAB24821.1; JOINED.
DR EMBL; S22203; AAB24821.1; JOINED.
DR EMBL; S22212; AAB24821.1; JOINED.
DR EMBL; S22213; AAB24821.1; JOINED.
DR EMBL; S22216; AAB24821.1; JOINED.
DR EMBL; S22219; AAB24821.1; JOINED.
DR EMBL; S22220; AAB24821.1; JOINED.
DR EMBL; S22221; AAB24821.1; JOINED.
DR EMBL; S22222; AAB24821.1; JOINED.
DR EMBL; S22226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RWHU1B.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genem; HGNC:6149; ITGAM.
DR MTM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VFWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

QY	1	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSLYOCYSTGSCPEI 60	1021	PFQIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGGAFVRSQTBET 1080
DB	17	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSLYOCYSTGSCPEI 76	1036	PFQIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGGAFVRSQTBET 1095
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 120	1081	KYEPPEVNPPLIYGVSSVGGLLLLALITAAALYKLGFFKRYKDMWSEGGPPGASBPQ 1137
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 136	1096	KYEPPEVNPPLIYGVSSVGGLLLLALITAAALYKLGFFKRYKDMWSEGGPPGASBPQ 1152
QY	121	FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLPFLSMQYSSEF 180	RESULT 2	
DB	137	FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLPFLSMQYSSEF 196	ITAM_MOUSE	
QY	181	RTHFTPEFQNNPNRPSLKPITQLGRTHATGVRKVIRELLNTNGARKNAFKILIVI 240	ID	ITAM_MOUSE
DB	197	RTHFTPEFQNNPNRPSLKPITQLGRTHATGVRKVIRELLNTNGARKNAFKILIVI 256	AC	P05555; STANDARD; PRT; 1153 AA.
QY	241	TGCEKFGDPLGYEDVPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDHFQVN 300	DT	01-NOV-1988 (Rel. 09, Created)
DB	257	TGCEKFGDPLGYEDVPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDHFQVN 316	DT	01-FEB-1991 (Rel. 17, Last sequence update)
QY	301	NFEALXTIQNLRKIFAIEGTQSGSSFEHMSQEGFSAAITSGNPLSTVGSYDWAG 360	DT	28-FEB-2003 (Rel. 41, Last annotation update)
DB	317	NFEALXTIQNLRKIFAIEGTQSGSSFEHMSQEGFSAAITSGNPLSTVGSYDWAG 376	DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI).
QY	361	GVFLYTSKSKSFTINTRVDSMDNDAYLGAAAILRNVRQSLVLCAPRYQHIGLVAMFR 420	GN	ITGAM.
DB	377	GVFLYTSKSKSFTINTRVDSMDNDAYLGAAAILRNVRQSLVLCAPRYQHIGLVAMFR 436	OS	Mus musculus (Mouse).
QY	421	QNTGMWSEANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQSVVCP 480	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB	437	QNTGMWSEANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQSVVCP 496	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
QY	481	PRGQARWQCDVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAVYLF 540	OX	NCBI_TaxID=10090;
DB	497	PRG-RARWQCDVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAVYLF 555	RN	[1] _TaxID=10090;
QY	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600	RP	SEQUENCE FROM N.A.
DB	556	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 615	RX	MEDLINE=88312584; PubMed=3044779;
QY	601	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660	RA	Pytela R.;
DB	616	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675	RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
QY	661	YDLALDSGRPHSRAVFNKSTRRTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720	RL	EMBO J. 7:1371-1378(1988).
DB	676	YDLALDSGRPHSRAVFNKSTRRTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 735	RN	[2]
QY	721	SLVGTPLSAFNGRLPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 780	RP	SEQUENCE OF 17-28.
DB	736	SLVGTPLSAFNGRLPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 795	RX	MEDLINE=85188776; PubMed=3887182;
QY	781	GPREFNVTVVRNDGSDSTVQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASSTEV 840	RA	Springer T.A., Teplow D.B., Dreyer W.J.;
DB	796	GPREFNVTVVRNDGSDSTVQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASSTEV 855	RT	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
QY	841	SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTFF 900	RL	Nature 314:540-542(1985).
DB	856	SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTFF 915	CC	ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS. EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
QY	901	QLELPVKYAVYVMVTHSGVSTKYLNFTASENTSRVNHQYQVSNLQORSLSPLSLVPLVP 960	CC	-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
DB	916	QLELPVKYAVYVMVTHSGVSTKYLNFTASENTSRVNHQYQVSNLQORSLSPLSLVPLVP 975	CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.
QY	961	RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020	CC	-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
DB	976	RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1035	CC	-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
			CC	-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
			CC	-!- SIMILARITY: Contains 1 VWFA domain.
			CC	-!- SIMILARITY: Contains 7 FG-GAP repeats.
			CC	-----
			CC	This SWISS-PROT entry is copyright. It is introduced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X07640; CAA30479.1; -
 DR EMBL: M14293; AAA39484.1; -
 DR PIR: S00551; S00551.
 DR HSSP: P11215; IABX.
 DR MGD; MGI:96607; Itgam.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP_3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Calcium; Repeat.
 FT SIGNAL 1 16
 FT CHAIN 17 1153
 FT DOMAIN 17 1105
 FT TRANSMEM 1106 1129
 FT DOMAIN 1130 1153
 FT REPEAT 31 84
 FT DOMAIN 164 350
 FT REPEAT 337 400
 FT REPEAT 401 452
 FT REPEAT 454 515
 FT REPEAT 517 575
 FT REPEAT 580 632
 FT CA_BIND 465 473
 FT CA_BIND 529 537
 FT CA_BIND 592 600
 FT SITE 1132 1136
 FT DISULFID 66 73
 FT DISULFID 105 123
 FT DISULFID 654 711
 FT DISULFID 770 776
 FT DISULFID 999 1023
 FT DISULFID 1028 1033
 FT CARBOHYD 58 58
 FT CARBOHYD 86 86
 FT CARBOHYD 391 391
 FT CARBOHYD 696 696
 FT CARBOHYD 734 734
 FT CARBOHYD 772 772
 FT CARBOHYD 801 801
 FT CARBOHYD 881 881
 FT CARBOHYD 907 907
 FT CARBOHYD 941 941
 FT CARBOHYD 980 980
 FT CARBOHYD 994 994
 FT CARBOHYD 1022 1022
 FT CARBOHYD 1045 1045
 FT CARBOHYD 1051 1051
 FT CARBOHYD 1076 1076
 SQ SEQUENCE 1153 AA; 127480 MW; 178DB98AECB0343 CRC64;

Query Match 76.2%; Score 4476; DB 1; Length 1153;
 Best Local Similarity 74.3%; Pred. No. 1e-294;
 Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;
 1 FNLDTENMTFQENARGQGSVQLOGSRVVVGAQBEIVAAVNRGSLYQCDYSTGSCPEI 60
 17 FNLDTENMTFQENAKGFGQNVVQGGTGVVVAAPQAKAVNQTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQOLLACGPTVHTQCSNTYVKGCLFLGFSNLRQPOK 120
 DB 77 PLQVPEAVNMSLGLSLAVTVPQOLLACGPTVHTQCKENTYVNGLYGLFSGNLLRPPO 136
 QY 121 FPEARLGCPEQSDSDIAFLIDGSGSIIPHDFRMKCFVSTVMEQLKSKTKLFLSLMQYSBF 180
 DB 137 FPEARLRECPQESDIVFLIDGSGSINNIDFQMKCFVSTVMEQFKSKTKLFLSLMQYSDF 196
 QY 181 RIHFTFKFQNNPNSLSVKPITOLLGRTHATTGVKRIRELLNTINGARKNAFKILVI 240
 DB 197 RIHFTFNDPKRNPSPRSHVSPKQLNGRTKTAGIRKVVRELFHKTNGARENAKILVI 256
 QY 241 TDGEKFGDPLGYEDVPIPEADREGVIRYVIGVGDAFRSEKSKROELANTIASKPRDHVQV 300
 DB 257 TDGEKFGDPLDYKDVPIPEADRAGVIRYVIGVGNFKNPQSRRELDITASKPAGERHVQVD 316
 QY 301 NFEALKTIONQUREKIFAIEGTQGTSSSFEHMSQEGFSAITSNGPLSTVGSVDWAG 360
 DB 317 NFEALNTIQLOEKIFAIEGTQGTSSFEHMSQEGFSAITSNGPLSGVSGSPDWAG 376
 QY 361 GVPLTYSKSKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLCAPYQHIGLVAMFR 420
 DB 377 GAFLTYSKOKVFINTRVDSMDNDAYLGAAVILNRVQSLVLCAPYQHIGLVAMFR 436
 QY 421 QNTGMESNANVKGTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYYEQTGCGQVSCPL 480
 DB 437 ENFGTWEPTSIGSQISGYFGASLCSDVDMDADGNTNLILIGAPHYYEKTGCGQVSCPL 496
 QY 481 PRQARWQCDVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAIGAPEEDNRGAYVLF 540
 DB 497 PRG-RARWQCEALLHGDQHPMGRFGAALTIVLDVNGDKLTDVAIGAPEEQENQGVYIF 555
 QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGODLTMDGLVDITVGAQGHVLLRSQ 600
 DB 556 YGASIASLSASHSHRIIGAHFSPGLQYFGOSLGGKDLTMDGUMDLAVGAQGHLLLRQA 615
 QY 601 PVLVRKAIMFENPREVARNVFECDQVVRKEAGEVRVCLVHVKSTRDLRREGQIOSVVT 660
 DB 616 PVLRLTEATWFEFSPKVARSVFACQEQVKNKADAGEVRVCLVRKNTKDLRREGDIQSTVT 675
 QY 661 YDLALSGRPHSPAVENETNSTRROTQVLGLTQTCETLKLQPLNCIEDPVSIVLRNLF 720
 DB 676 YDLALDPVRSIRAFEDTKNTRRTQVFLGQKCEKTLKILPDCVDSVSIILRLNY 735
 QY 721 SLVGTPLSARGNLRPVLAEDAOELFTALPFEKNCGNDNICODDLSITFSFMSLDCLVWG 780
 DB 736 TLVGEPLRSFNLRPVLAEDAQORFFTFAMPFPEKNCGNDISICQDDLSITMSAGLDTLVWG 795
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWRL-ACESASSTE 839
 DB 796 GPQDFNMSVTLRNDGEDSYGTQVTVYPSGLSVRKDSASQNPITKXPMFKVPAESSSSSE 855
 QY 840 VSGALKSTGCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSNNPRNTKTE 899
 DB 856 GHGALKSTTNINHPFPANSEVTFNVTFDVDSHSGFNKLLKAIIVASENNMSRTHKTK 915
 QY 900 FQLELPVKYAVYVVTSHGVSTKYLAFTASENTSRVMOHQYOVNLSGORSPLSLVPLVP 959
 DB 916 FQLELPVKYAIYVVTSDSSIRYLAFTASEMTSKVIOHQYFNNLQGRSLPVSVPWIP 975
 QY 960 VRLNQTVMDRPOVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNVCSIAVCQICDI 1019
 DB 976 VQINNVTVMDHPQVIFSQNLSSACHTEQKSPHSHNFRDQLERTPVLNCSVAVCKRIOC 1035
 QY 1020 PFFGIGQEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQTE 1079
 DB 1036 PSFNTQEIFNVTLKGNLSPDWYIKTSHGHLLVSSSTEILFNDVSFAFALLPGQSVRSKTE 1095
 QY 1080 TKVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYOKDMMSGEGPGABEQ 1137
 DB 1096 TKVEPEVNPPLIIVGSSIGGLVLLALITAGLYKLGFFKRYOKDMMSAEPADAPQ 1153

RESULT 3
 ID TAX_HUMAN STANDARD; PRT; 1163 AA.
 AC P20702;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
 GN alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
 DN ITGAX OR CD11C.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8816645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of
 a leukocyte adhesion glycoprotein, p150,95";
 RL EMBO J. 6:4023-4028(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte
 p150,95 molecule";
 RL J. Biol. Chem. 265:2782-2788(1990).
 [3]
 RP ERRATUM.
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751(1990).
 [4]
 RP SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 and p150,95 leukocyte adhesion proteins";
 RL J. Immunol. 138:2381-2383(1987).
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
 RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
 INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
 IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
 ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VFMA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M81695; AAA59180.1; -;
 DR EMBL; Y00093; CAA68283.1; -;
 DR EMBL; M29165; -; NOT ANNOTATED CDS.
 DR EMBL; M29487; AAA51620.1; ALT_SEQ.
 DR EMBL; M29482; AAA51620.1; JOINED.
 DR EMBL; M29483; AAA51620.1; JOINED.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29485; AAA51620.1; JOINED.

DR EMBL; M29486; AAA51620.1; JOINED.
 DR PIR; A36584; RWHUIC.
 DR PDB; 1N3Y; 18-FEB-03.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1163 INTEGRIN ALPHA-X.
 FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1108 1128 POTENTIAL.
 FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 165 351 VWFA.
 FT REPEAT ? ? FG-GAP 3.
 FT REPEAT 402 453 FG-GAP 4.
 FT REPEAT 455 517 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA BIND 466 474 POTENTIAL.
 FT CA BIND 530 538 POTENTIAL.
 FT CA BIND 593 601 POTENTIAL.
 FT SITE 1131 1135 GFPR MOTIF.
 FT DISULFID 69 76 BY SIMILARITY.
 FT DISULFID 108 126 BY SIMILARITY.
 FT DISULFID 655 712 BY SIMILARITY.
 FT DISULFID 771 777 BY SIMILARITY.
 FT DISULFID 848 863 BY SIMILARITY.
 FT DISULFID 998 1022 BY SIMILARITY.
 FT DISULFID 1027 1032 BY SIMILARITY.
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 490 490 G -> A (IN REF. 2).
 FT CONFLICT 756 756 L -> D (IN REF. 2).
 SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;

Query Match 59.1%; Score 3473; DB 1; Length 1163;
 Best Local Similarity 61.3%; Pred. No. 7.8e-227;
 Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;
 Qy 1 FNLDTENAMTFQBNARFGSGVVOLOGSRVVVVGAPQEIIVANQSGSYQCDYTGSCPEI 60
 Db 20 FNLDTEELTAFRVDSAGFSGSVVQYANSVVVGAPQKITAANQTGGYQCGYSTGACEPI 79
 Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLIACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
 Db 80 GLQVPVEAVNMSLGLSLASTTSPQLIACGPTVHCEGRNMYLTGLCLLGLPT--QLTOR 137
 Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKSKTLFLSLMQYSEEF 180
 Db 138 LPVSRQECPRQEQDIVFLIDGSGSISRRNFATMWNFVRAVISQFRPSTQFLMQFSNKF 197

DB 1091 MVLEDEVYNAIPIMSSVGALLALITATLYKLGFFRHYKEMLED 11139

RESULT 5

ITAL_HUMAN STANDARD; PRT; 1170 AA.

AC P20701; O43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).

GN ITGAL OR CD11A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=2537322;

RA Larson R.S., Corbi A.L., Berman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";

RT J. Cell Biol. 108:703-712(1989).

RL [2]

RN SEQUENCE FROM N.A. (ISOFORM 2)

RX MEDLINE=99425270; PubMed=10493829;

RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 1p and 16q.";

RT Genomics 60:295-308(1999).

RL [3]

RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Qu A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";

RT Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).

RL [4]

RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398692; PubMed=8805579;

RA Qu A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";

RT Structure 4:931-942(1996).

RL [5]

RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=99425288; PubMed=10493852;

RA Kallen J., Weizenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain.";

RT J. Mol. Biol. 292:1-9(1999).

RL CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2, ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES AND MONOCYTES.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=P20701-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P20701-2; Sequence=VSP_002738;

CC Note=No experimental confirmation available;

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry; WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Y00796; CAA68747.1; -.

DR EMBL; AC002310; AAC31672.1; -.

DR PIR; S03308; S03308.

DR PDB; 1LFA; 29-JAN-96.

DR PDB; 1ZON; 07-DEC-96.

DR PDB; 1ZOO; 07-DEC-96.

DR PDB; 1ZOP; 07-DEC-96.

DR PDB; 1COP; 07-AUG-00.

DR PDB; 1DQ; 03-FEB-00.

DR PDB; 1MQ8; 14-JAN-03.

DR PDB; 1MQ9; 14-JAN-03.

DR PDB; 1MQA; 14-JAN-03.

DR MIM; 153370; -.

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0006928; P:cell motility; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS0234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; 3D-structure; Magnesium; Calcium; Repeat;

KW Alternative splicing.

FT SIGNAL 1 25

FT CHAIN 26 1170 INTEGRIN ALPHA-L.

FT DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1089 1112 POTENTIAL.

FT DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

FT REPEAT 42 91 FG-GAP 1.

FT REPEAT 92 149 FG-GAP 2.

FT DOMAIN 170 349 VWFA.

FT REPEAT 349 401 FG-GAP 3.

FT REPEAT 401 455 FG-GAP 4.

FT REPEAT 457 516 FG-GAP 5.

FT REPEAT 518 575 FG-GAP 6.

FT REPEAT 578 630 FG-GAP 7.

FT CA_BIND 468 476 POTENTIAL.

FT CA_BIND 530 538 POTENTIAL.

FT CA_BIND 590 598 POTENTIAL.

FT SITE 1115 1119 GFPR MOTIF.

FT DISULFID 73 80 BY SIMILARITY.

FT DISULFID 111 129 BY SIMILARITY.

FT DISULFID 653 707 BY SIMILARITY.

FT DISULFID 771 777 BY SIMILARITY.

FT DISULFID 845 861 BY SIMILARITY.

FT DISULFID 998 1013 BY SIMILARITY.

FT DISULFID 1021 1052 BY SIMILARITY.

FT CARBOHYD 65 N-LINKED (GLCNAC...) (POTENTIAL).

NCBI TaxID=10090;
(1) _
RN SEQUENCE FROM N.A. PubMed=2051027;
RP MEDLINE=91268576; PubMed=2051027;
RX Kaufmann Y., Tseng E., Springer T.A.;
RA "Cloning of the murine lymphocyte function-associated molecule-1
RT alpha-subunit and its expression in COS cells.";
RN J. Immunol. 147:369-374(1991).
(2)
RN SEQUENCE OF 24-42.
RP MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC !- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
CC LEUKOCYTES RECRUITMENT.
CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC ASSOCIATES WITH BETA-2.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- TISSUE SPECIFICITY: LEUKOCYTES.
CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC !- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC !- SIMILARITY: Contains 1 VWFA domain.
CC !- SIMILARITY: Contains 1 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60778; AAA39426.1; -
DR PIR; I56126; I56126.
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR PRINTS; PR01185; VWFA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium;
KW Repeat.
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT CA_BIND 528 536

	CA_BIND	588	596	POTENTIAL.
FT SITE	1111	1115		GFPR MOTIF.
FT DISULFID	70			BY SIMILARITY.
FT DISULFID	108	126		BY SIMILARITY.
FT DISULFID	147	199		BY SIMILARITY.
FT DISULFID	651	705		BY SIMILARITY.
FT DISULFID	767	773		BY SIMILARITY.
FT DISULFID	840	856		BY SIMILARITY.
FT DISULFID	993	1009		BY SIMILARITY.
FT DISULFID	1017	1048		BY SIMILARITY.
FT CARBOHYD	86	86		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	185	185		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	270	270		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	444	444		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	668	668		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	696	696		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	724	724		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	728	728		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	776	776		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	857	857		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	880	880		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	890	890		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	899	899		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	927	927		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1056	1056		N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE	1163 AA; 128343 MW; A7A3078489E8232F CRC64;			

Query Match 26.1%; Score 1533.5; DB 1; Length 1163;
Best Local Similarity 34.0%; Pred. No. 1.4e-95;
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;

QY	1	FNLDTENAMTFOENA-RFGQSVVQLQSGRVVGAQPBIVAAQORSGSLYQCDYSTGSC	59
DB	24	YNLDTPTQSLAQAGRHFGYQVLIQEDG-VVVGAPGE---GDNTGGLYHCRTSSEFCQ	79
QY	60	IRLQVPEAVNMSGLSLAATSPQQLACGPTVHQCSTNTYKGLCFLGSNLRQPPQ	119
DB	80	VSLH-GSNHTSKYLGMTLATDAKGSLLACDPLGSLRTCDQNTYLSGLCYLPQSL	138
QY	120	KPPALRCPCDSDIAFLIDSGSIIPHDFRRMKFVSTVMEOLKSKTKTFLSLMOYSEE	179
DB	139	QNRPAYQECMKGVLDVFLFQSGSLDRKDFEKMVDWKLNSNTSYQFAAVQFSTD	198
QY	180	FRIHFTKEF-QNNPNRSLVKPITQLLGRTHATTGVKRVIRELLNITNGARKNAFKLI	238
DB	199	CRTEFTLDYVQKNKPNPDVLLGSVQPMFLLNTFRAINVVAHVFPKESGARPDATKLV	258
QY	239	VITGKEGDFLGYEDVPIEADREG-----VIRYVIGVGDAFSEKSRQELNTIASKP	291
DB	259	IIIDG-----EASDKGNISAAHDITRYIIIGKHGFVSQKQTLHFASEP	304
QY	292	PRDHVQVNFPEALKTIQNLREKIFAETGTQSGSSSFEHEMSQEGFSAITNSGPLLS	351
DB	305	VEEFVKIIDTPEKLDLTDLQRRYIALEGTRQDLTSFNMLSSSGISADLSKHAVVG	364
QY	352	TVGSDYDAGGVF-LYTSKEKSTFTINMTRVDSMDNDAYLGAAAA-IILNRVOSLVLAGP	409
DB	365	AVGAKDWAGGFLDLREDLQAGTFVQGEPLTSDVRGGYLGTYVAMWTSRSSRPLLAAGAP	424
QY	410	YOHIGLVAMFR--QNTGMWESNANVKGTOIGAYFASLCSVDVDSNGSTDLVLICAPHY	467
DB	425	YQHVQVLLFOAPEAGGRWNQTKIEGTQIGSYFGELCSVDLDQDGAELLIGAPLFF	484
QY	468	EQTRGGOVSVCPLPRGORARWOCDAVLVGEQOPMGREFGAALTVLGDVNGDKLTDAVGA	527
DB	485	GEQGRGVFTY---QRRQSLFEMVSELOQDPGYPLGRFGAAITALTINDGRDLTDAVGA	541
QY	528	PGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKSLPRLOYFQOQSLGGQDLTMDGLVDLT	587
DB	542	PLEE--QGAIVIFNCKPG-GLSPQSPQRIQGAQVFPGRWFGRSIHGVKDLGGDLADVV	598
QY	588	VGAQGHVLLRSQPVLRVKAIMEFNPVARNVFCNDQOVVKGEAG-EVRVCLHVQKST	646

Db 599 VGEGRVVVLSRRPVDVVTLSFSPBEPVHEVECSYAREEQKHGKVKACPRIKPLT 659
 Qy 647 RDRLEGGIOQVWYDIALDSGRPHSRAVFNETKXSTRQVTLGLTQTCTELKLQLPNC 706
 Db 659 PQ--FQGRLLANLSTYQLQDGHMRSGCLPFDGSHSLSGNTSITP-DKSCLDFFHFHPIC 715
 Qy 707 IEDVPSIVLRNLSLV--CTPLSAGFN-LRPVLAEADAQLFTALPPFKKNCNDNICQ 762
 Db 716 IQDLISPTNVLSNLSLEEGTPTDQKRAMQPIRLPSIHTV-TKEIPPFKNCGEDKKCE 774
 Qy 763 DDLSTISFMSLDCLVVGGP-----REFNVTVVRNDGDSYRTQVTFPPFLDLSYRKV 816
 Db 775 ANLTSSPARS-----GURLMSSASLAVETLNSGSDAYWRLDLPFRGLSPKRV 827
 Qy 817 STLQQRQSRWRLACSSASTVSGAL-KSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
 Db 828 EMLO---PHSRMPVSCEEL--TEGSULLTKLKNVSSPIFKAGQEVSLQVMFNTLNS 882
 Qy 876 LGNKLKLLKANTYSEN-NMPTNKTEFQLELPVKVAVYVWVTSVGHVSKYLNFTASENTR 934
 Db 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLYPVNLTKEQENSTLYISFTPKGPKTQ 942
 Qy 935 VMHQYQVSNLQGRSLPSLVLPVRLNQTVIDRPO-----VTFSENLS--TCHTK 986
 Db 943 QVQHVYQV-----RIQPSAYDHNMT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996
 Qy 987 E-RLPSSHDFLAELRKAPVNCSTAVCORIQDIPFFQIOBEFNATLKNLSFDWYIKTS 1045
 Db 997 DLKPSSE---AEQCLPGV-----QRCPIVF---RWEILIQVTGVELSKEIKAS 1042
 Qy 1046 HNHLIVSTABILFNDVSFTLLPQOGAFVRSQTEKVEPFPVNPPLPIVGVSSVGGLLLL 1105
 Db 1043 -STLSCLSSLSVSNSSKHFLYGSKA-SEAQLVKVDLIEKEMLRVYVLSGIGGLVLL 1100
 Qy 1106 ALITAAALKLGFRRQYKDMV-SEGGPGGAP 1136
 Db 1101 FLIFLALYKGVFFKRNKLEKMEADGGVNGSP 1132

RESULT 7
 ITAE_MOUSE STANDARD; PRT: 1167 AA.
 AC Q60677;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-E precursor (Integrin alpha M290).
 GN ITGAE;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=AKR;
 RX MEDLINE=95187992; PubMed=7882170;
 RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
 RA Kilshaw P.J., Wells J.H.;
 RT "Murine M290 integrin expression modulated by mast cell activation.";
 RL Immunity 1:393-403(1994).
 CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
 CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
 CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
 CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
 CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
 CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
 CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
 CC INTESTINAL WALL.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
 CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; U12236; AAC52142.1; -
 CC HSP; P11215; 1A8X.
 CC MGD; MGI:1298377; Itgae.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 3.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS02234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Signal; Repeat; Magnesium;
 CC Calcium.
 CC SIGNAL 1 19 BY SIMILARITY.
 CC CHAIN 20 1167 INTEGRIN ALPHA-E.
 CC CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
 CC CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
 CC DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 1115 1137 POTENTIAL.
 CC DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
 CC REPEAT ? ? FG-GAP 1.
 CC REPEAT ? ? FG-GAP 2.
 CC DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
 CC DOMAIN 193 384 VWFA.
 CC REPEAT 449 ? FG-GAP 3.
 CC REPEAT 503 564 FG-GAP 4.
 CC REPEAT 566 631 FG-GAP 5.
 CC REPEAT 634 686 FG-GAP 6.
 CC CA_BIND 514 522 FG-GAP 7.
 CC CA_BIND 578 586 POTENTIAL.
 CC DOMAIN 185 191 POTENTIAL.
 CC SITE 1140 1144 GLU-RICH (ACIDIC).
 CC DISULFID 72 83 GFPR MOTIF.
 CC DISULFID 130 164 BY SIMILARITY.
 CC DISULFID 698 754 BY SIMILARITY.
 CC DISULFID 814 820 BY SIMILARITY.
 CC DISULFID 884 898 BY SIMILARITY.
 CC DISULFID 998 1023 BY SIMILARITY.
 CC DISULFID 1031 1047 BY SIMILARITY.
 CC CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 341 341 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 718 718 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 829 829 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 846 846 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 925 925 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;

Query Match 19.7%; Score 1157.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 4.1e-70;
Matches 358; Conservative 211; Mismatches 459; Indels 205; Gaps 43;

QY 1 ENLDTEA--MTFOENARGFGSGVVOLOGSRVVVVGAPQBIVAANORGS-----LYQCDY 52
DB 20 FNVVDVMAVWVTLQPGAPAVLSLLHLDPSN-----NQTLVARSSNRNTAALYRCAL 74
QY 53 STGSCSEPIRLQVPEAVNMSLSLSLAAT--SPPOLLAC-GPTVHQTCSENTYVKGCLFL 109
DB 75 SI-SPDEIACQ-PVEHICMPKRGYQGVTLVGNHGVLCIQVQARKFRSLNSELGTACSL 132
QY 110 FGSNLROQKQKPEALRG-----C-----PQS 131
DB 133 LTPNLDLQAQVPSDLGFLDPCAHVDGSDYCRSKGSGTGEEKSARRRRRTVEEDED 192
QY 132 DSDIAPLIDSGSIIIPHDFRMKEFVSTVMEQL--KSKTSLFSLMOYSEEPRIHFTKFEF 189
DB 193 GTEIAIVLDGSGSIPGSDFOKKNFISTWNRNFEYKCFECFNALVQYGAIVQTEFDLQES 252
QY 190 QNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVITDGBKFGDP 249
DB 253 ROLNASLAKVQSIQVQVKEVTKTASAMQVLDNIFIPSRGSRKALKVMVLDGDFGDP 312
QY 250 LGYEDVIPADREGVIRYVIGVDAPFRSEKSEKQELNTIASKPPRDHVFQVNNFEALKTIQ 309
DB 313 LNLTTVINSPKMGQVVRFAIGVDRFKNNNTYRELKLIASDPKEATFKVTNYNSALDGLL 372
QY 310 NQREKIPALEGTQTCSSSEFHEMSEGEFSAAITNSGP-LISTVGSYDWAQGVLY-TS 367
DB 373 SKLQQRIVHMEGT--VGDALQYQLAQTFGSAQILDKGVLGTGAFNWSGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGVAATILNRVQSLVGLGAPRYQHIGLVAMFRONTQM 425
DB 430 NRGCCFLNQTKEDSRTQVSYLYGLYSLAVLHKAHGISYVAGAPRHLKGAIVELRDEDR- 488
QY 426 WESNA---NVKGTQIGAYFGASICSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCPLPR 482
DB 489 -EEDAFVRRTEGRQMGSYFGSLCPVDIDMDGTDFLLVAAPFYHIRGEEGRVYVQVPE 547
QY 483 GQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAGP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHFGLTNSRFGFAVAAGVDINQDKFTDVAIGAPLEGAGDGASYS 606
QY 537 VYLFHGTSGSGISPSQRSIAGSKLSPRLOYFGSLSGGQDLTMDGLVLDLTGVAQGHVLL 596
DB 607 VYLYNGHSG-GLYDPSQDIASSVAGSLHYFGMSVSGGLDFNGDGLADITVGSRSADV 665
QY 597 LRSQPLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMFTFP-----DALPMVFIGKM--DVNLCEFDSSVASEPGLREM 715
QY 654 QIGSVVTYDLALSGRPHRAVNEKYNSTRQTQVLTQTC-----696
DB 716 FLNFTVDVVDV-----TKQRQLQCEDSSGQCLRKWNKNGSFLCEHFWLI 760
QY 697 ETILKLQLPNCIEDVPSPVILNLFSLVGTPLSAFGLNR-----PVLAEADAQRLFTALF--P 750
DB 761 STEEL-----CEEDCFSNITIKTYE-----FQISGGRDYPNTL--DHYKPSAIFQLP 809
QY 751 FERKCGNDNTCODLSITFSFMSLDCLVVGPGPEFNVTVTRVNDGSDSYRTQVTFPPDL 810
DB 810 YEKDCKNKVFCIAEIQLTTN-ISQQLVVGVTKEVTMNIISLTNSGSDSYMTNMAVYPRN 868
QY 811 LSVKRVSTLQNRQSRWRACSSASTEVSGALKTSKCSINHIPIIPENSEVTNFTFDV 870
DB 869 LQFKKI-----QKPSFDPVQCDPKPV--ASVLMNCKIGHPIIL-KRSSVNVSVTWQL 918
QY 871 DSKASUGNKLKLLKANVTSENNMPTNKTETEQLELPVKYAVYVMVTVSHGVSTKVLNFTASE 930

DB 919 EESVFPNRTADITVTSNSNEKSLARETR---SLQFRHAFIAVLRS--PSVMYMN--TSQ 971
QY 931 NTSRVMOHQVQSVNLGQSRSLPISLVFLVPVRLNQTIVWDRPQVTFSENLSST-----CHT 985
DB 972 SPDSHKKEFFNVHGENLFGAVFQLOICVPIKLQDF-----QIVRVKNLTKTQDHTSCTQ 1025
QY 986 KERLPSSHSDFLAELRKAPVNCVSIACVQRIQCDIPFFGIOBEFNATLKNLSFDWYIKTS 1045
DB 1026 SQEPACGSDPQVHKWHSVVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLIIVSTA-----EILFNDSSVFTLLPGGAFVRSQTEKVEFPF-----EVPNPLPLIV 1095
DB 1066 HTKOLLRDVSELPLIGEISFNKSLYEGLNAB-----NHRTKITVIFLKEETSRLPLII 1119
QY 1096 GSSVGGGLLLLIITAALYKLGFFKQYQKMMSE 1128
DB 1120 GSSIGGLLVLVIIAIFKCGFFKRYQQLNLE 1152

RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-1EL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RC MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RC MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shetlersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
RA "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

Qy 107 CFFGSLNRQOPQKPPBALRGCPQEDSDIAFLIDGSGIIPHDPFRMKFVSTVMEQLK- 165
Db 120 CSDVSPFTQVNVNAP--VOECSTQ-LDIIVLDGNSIYPWD--SVTAFNLLDKEMDI 174
Qy 166 -KSKTLESNQSEEFRIHTFKFQNNPNRSLVKBITQLLGR-THTATGVKRVIRELL 223
Db 175 GPKQTQVGIYOYGVNTHFENLNKYSSTEEVLAACKIVQGGRTWTALGTDTARKEAF 234
Qy 224 NITGARKNAFKILIVITDCEKFGDPLGYEDVPEADREGVIRVIVGVDAFR-----SE 278
Db 235 TEARGARGVKKVWIVITDSEH-DNHLKKVITQDCEDENIQRPSIAILGYSYRGNLSTE 293
Qy 279 KSRQELNTIASKPRDRHVQNNPEALKTIONQLREKIFALEGTQTGSSSFHEMSQEG 338
Db 294 KFBVEIKSIASEPTEKHFNFVSDLEALVTIVKTLGERIFALEATADQSAASFEMWSQTG 353
Qy 339 FSAITSGPLSLTSGVDHAGGVFLYTSKE-----KSTF-INWTRVDSMDNAYILGYAA 392
Db 354 FSAHYSQDWMLGAVGAYDNGVTVMQASQIIIPRNTTFNVSTKQNEFL-ASYLGYTV 412
Qy 393 AILLRNRVQSL-VLGAPRYOHIGLVAMFRQNTGMESNANVKGTQIGAYFGASLCSVDVD 451
Db 413 NSATASSGDVLYAGQPRYHNTGVIIYRMEDGNKILQTLGSEQIGSYFGSILTTTDD 472
Qy 452 SNGSTDLLVLGAPHY-----YEQTR-CGQVSVCPPLPRGQARMQCDAVLY 495
Db 473 KDSNTDILLGAPWYMGTEKEEQGVVYVALNQTREYQMSLEPIKQTCSSRQHNSCTT 532
Qy 496 GEQOPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNDRGAVYLFHGTSGSISPSHSQ 554
Db 533 ENKNEPCGARGFTGIAAIAVKOLNLDGPNFIDVIGAPLEDHGGVAYVYHG-SGKTIRKYEAY 591
Qy 555 RIAGSKLSPLOYFGQSLGQDLTMDGLDVLTVGAQGHVLLRSQVPLKAVKAIMFNPR 614
Db 592 RIPSGGDKTLKFPQSIHGMNDLNGDGLTDTVTIGLGGAALEFWSRDVAVVKTMMFEPN 651
Qy 615 EVARNVEFCNDQVVKGEAG--EVRVCLHVQ-KSTRDLREGQTQSVVTVYDLALDGRPH 671
Db 652 KVNQKKNCH--MEGETVCINATVCFVKLKSKEDTIYEAQLQ---YRVTLDSLRLQ 704
Qy 672 SRAVFNET-----KNSTRQQTQVLGTQTCETLKLQPLNCIEDPVPVILRNFSLVGT 725
Db 705 SRSPFSGTQBRKVQRNITVAKSEC-----TKHSPYMLDKHDFQDSVR---ITLDFNLT-D 755
Qy 726 PLSAFGNLRPVLAEADQLFTALPPEKNCNDNICODDLISITFSPSLDCLVVGPRE- 784
Db 756 PENG-----PVLDDSLNSVHEYIPFAKDCGNKEKICISLSLHVATTEKOLLIVRSQNDK 810
Qy 785 FNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEVSGAL 844
Db 811 FNVSLTVKNTKDSAYNTRTIVHYSPLNLFVSGIEAIQKD-----SCSN----- 853
Qy 845 KSTSCSINHPIFPENSEVTNITPDVDSKASLGN-KULLKANVTSENNMPTKNTPEQLE 903
Db 854 HNITKGVGPPLRRGEMVTIKLPQNTSYLMENVTVIYLSATSDSEPPETLSNVVNIS 913
Qy 904 LPVKYAYVMVTVSHGVSTKYNFTASENTSRVMOHQYQVSN-----LGQRS-----L 950
Db 914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPMP 972
Qy 951 PISLVF-----LVPVRLNQTIVWDRPQVTFNSLSTCHTK-----RLPS 991
Db 973 KLSISFPMNTSGVPLVYPTGLSS-----SENANCRPHIFEDPSINSKGKMTT 1021
Qy 992 HSDFLAELKAPVNVNCSIAVCRIQCDIPFFGIOE-----EFNATLK 1033
Db 1022 STD---HLKRGITLDCNTCKFATITCNLTSSDISQVNVSLIMKPTFIKVSFSLNLTIR 1078
Qy 1034 GNLSFDWIKYKSHNHLIVSTAELFNDSVFTLLPGOGAFVRSQTETKVFEPFVNPPL 1093
Db 1079 GEL-----RSENASIVLSSEN-----QKRELAIQISKDGLPGRVPL 1114
Qy 1094 --IVGSSVGGLLLALLALITAAALYKLGFFKQYKDMWSE 1128

Db 1115 WILLSAFAGLLMLLLALWKIGFFKRPLKKOMEK 1151
RESULT 10
ITAH HUMAN STANDARD; PRT: 1189 AA.
ID ID ITAH HUMAN STANDARD; PRT: 1189 AA.
AC Q9UKX5; Q9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITG11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA Wang S.-X., Morris C.M., Kriessansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
RL human integrin alpha11 subunit (ITG11).";
RL Genomics 60:179-187(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kueche-Gullberg M., Sejersten T., Gullberg D.;
RT "CDNA Cloning and Chromosomal Localization of Human alpha(11)
RL integrin. A collagen-binding, i domain-containing, beta(1)-associated
RL integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
RP SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Bativill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC -!- ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 EG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF109681; AA01258.1; -
CC EMBL; AF137378; AAD51919.2; -
CC EMBL; AL359064; CAB94392.1; -
CC HSSP; P17301; IAOX.
CC Genew; HGNC:6136; ITG11.
CC MIM; 604789; -
CC GO; GO:0008305; C:integrin complex; TAS.
DR

DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0005518; F:collagen binding activity; TAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO; GO:0007517; P:muscle development; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; WVFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; vwa; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS0234; WVF_A.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 1189
 FT DOMAIN 23 1142
 FT TRANSMEM 1143 1165
 FT DOMAIN 1166 1189
 FT REPEAT 38 94
 FT REPEAT 102 163
 FT DOMAIN 167 345
 FT REPEAT 359 420
 FT REPEAT 422 475
 FT REPEAT 477 537
 FT REPEAT 539 598
 FT REPEAT 601 653
 FT DOMAIN 1154 1162
 FT DOMAIN 1174 1177
 FT CA_BIND 488 496
 FT CA_BIND 511 559
 FT CA_BIND 613 621
 FT DISULFID 76 83
 FT DISULFID 121 139
 FT DISULFID 129 159
 FT DISULFID 659 668
 FT DISULFID 674 729
 FT DISULFID 781 787
 FT DISULFID 881 893
 FT CARBOHYD 82 85
 FT CARBOHYD 95 95
 FT CARBOHYD 291 311
 FT CARBOHYD 331 331
 FT CARBOHYD 358 358
 FT CARBOHYD 449 449
 FT CARBOHYD 462 462
 FT CARBOHYD 528 528
 FT CARBOHYD 642 642
 FT CARBOHYD 694 694
 FT CARBOHYD 857 857
 FT CARBOHYD 894 894
 FT CARBOHYD 973 973
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1040 1040
 FT VARIANT 433 433
 FT VARIANT 524 524
 FT VARIANT 972 972
 FT VARIANT 1003 1003
 FT VARIANT 1030 1030
 FT VARIANT 1094 1094
 FT SEQUENCE 1189 AA; 133609 MW; 60303C08A44CD52 CRC64;

Query Match

18.6%;

Best Local Similarity 28.3%;

Pred. No. 9e-66;

Matches 349; Conservative 214; Mismatches 502; Indels 167; Gaps 47;

1 FNLDTNAMTFQENARG-FGQSVVQ--LOGSR-VVVGAPQBIIVAANQGRSLYQCDYSTGS 56
 23 FNDTRKPRVPGSRTAFFGYTQQHDSGNKMLVVGAPLETNGYQKTGDVYKCPVIHGN 82
 57 CEPRL-----QVPVEAVNMSLGLSLAATSPQLLACGTPVHQTCSNTYVYKGLFLP 110
 83 CTKLNLGRVTLNVSERKDNRLGLSLATNPKNPSFLACSPLSMSHECGSSYTTGMSRV 142
 111 GSNLRQQQPKFPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTWMEQ--LKSK 168
 143 NSNFRFSKTVAP-ALQRC-QTYMDIVLDGNSIYP--WVEVHFNLILKKFYIGPGQ 198
 169 TFLSLMQSYBEFRIHFTKFEQNNPNRSLVKPTQLLGR-KTHPTATGVKRVIRELLNITN 227
 199 IQGVVQYGEDVHFEFLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQ--K 256
 228 GARKNAKILIVTDGEKFGDPLYEDVIPADREGVIRVIGV-----GDAFRSEKSRQ 282
 257 GGRGAKKVMIVITDGESHOSP-DLEKVIQOESRDNVTYAVAVLGYNNRRGINPETFLN 315
 283 ELNTIASKPPREDHVQVNNFEALKTIONLREKIFAIEGTQGTSSSSFEHEMSGEFSAA 342
 316 EIKYIASDPDDKHFFNTDEALKDIDVLDGDRIFSLGNTK-NETSFGLMSQTGFSH 374
 343 ITSNGPLLSVGSYDWAGGVFLYTSKEK-----STFINMTRVDSMDNDAYLGAAAIILR 397
 375 VVEDGVLLGAVGAYDMNGAVLKETSGAKVIPRESYLKEFPEELKNHGAYLGYTVTSVVS 434
 398 NRV-OSLVLGAPRYQHIGLVAMP-RONTGMWESNANVKGTOIGAYFGASCLSCVDVDSNGS 455
 435 SROGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMRGQIQSGYFSGEITSVDIGDGV 494
 456 TDVLIGAPHVYEQTR-GGQVSVCPPLRGQARQOCDAVLVGEQGPNGRGAALTLVGD 514
 495 TDVLVGAPMYFNREGRGKVITYEL---RQNRVYNGTLKDSHSYQNARFGSSIASVRD 551
 515 VNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFQGSLSG 574
 552 LNQDSYNDVVVGAPLEDNHAGAIYIFHGFRGS-ILKTPKQRIITASELATGLQFGCSIHG 610
 575 QDLTMOGLVDLTGVAQGHVLLRSQPVLRVKATMEFNPREVARNVF--ECNDQVVVKGE 632
 611 QLDLNEGLIDLVAGALGNVILMSRPVQVNASLHFEPSKI--NIFHRDC-----KR 661
 633 AGEVRVCL-----HVOKSTRDLREGIQSVVTDYLDLSDGRPHSAFVNET 679
 662 SGRDATCLAAFLCFTPIPLAPHFQTTVG-----IRYNATMDERRITPRAHLDEG 711
 680 KNS-TRRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNPSLVGTPLSAFGLNLRPLVA 738
 712 GDRFTNRAVLSSGOELCERINFHVLDTADYVVPVTFSEVSELEDP-----DHGPMLD 764
 739 EDAQRLLFTALPFPEKKNQNDNICQDDL-----SITSPF 771
 765 DGMPTTLRVSPVFWNGCNEDEHCVPLDVLDAESDLPTAMEYQORVLRKPAQDCSAYTLSP 824
 772 MSLDCLVVGGRPFNVTVTVENDGDSYRTQVTFPEPLDLSYRKVSTLQNRQSRWRLA 831
 825 DTTVFIIESTQRVAVEATLENRGENAYSTVLNISQSANLQF--ASLQKEDSDGS--IE 880
 832 CESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENN 891
 881 CVNEER-----RLQKQVCNVSPFPRAKAKAVAFRLDPEF-SKSIPLHLELELAAGSDSN 934
 892 MPRTKNTE--FOLELPVYKAVVMVTVSHGVSTKY---LNFTAS--ENTSRVWQHOYQVSN 944
 935 ERDSTKEDNVAPLRPHLYEADVLFTRSSLSLHYEVKLNSSLERYDVGIPGPFSCIFRQN 994
 945 LQO----RSLPLSLVPLVPLVRLNQTVWDRPQVTFSENLSSTC-----HTKERLPSSHDFLA 997
 995 LGLFPIHGIMMKITIPATRSNRLKLLKLRDLFT-DEVANTSCNINWNTSTYRTPVE--E 1051

Db 370 VGFSAEYSPQNNILMGAVGAYDWSGTVVQKTPHGHILFISKQAFQILQDRNHSYLGYS 429
 Qy 392 AAILLRNRVQSLVLCAPRYOHIGLVAMFRQNTGMESNANV-----KGTOIGAVFGASL 445
 Db 430 VASISTGNSVHFVAGAPRANYTGQIYLSVN-----ENGNTVVIOSQGDQIGSYFGSVL 484
 Qy 446 CSVDVDSNGSTDLVLIGAPHYEQTR--GGQVSVCPPLPRGQPARQWCDAVLYGEGCQPMW 503
 Db 485 CAVONKOTITDVLVAGPMYNDLKKEGRVYLFITIKG-ILNWH--QFLEGNPLENA 541
 Qy 504 RFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVLYFHTSGSGISPSHSORIAS--KL 561
 Db 542 RFGSAIALSDINDMGDFNDVIGSPLENGNSGAVIYNGHEGM-IRLRSYKILGSDRAF 600
 Qy 562 SPRLOYGOSLGGODLTWDGLVDLTGGAQGHVLLLRQPVLRVKAIMEFNPREVARNVF 621
 Db 601 SSHLOYFGRSLDGYGDLNGSDITDVSAGFGVQVQLWSQSIADSVDSFPPKIL--TIL 658
 Qy 622 ECNDQVVGKEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDIALD-----SGRPHSRAVFN 677
 Db 659 NKNAEI-----KLKCF-----SAKFRPTNQNNQVAIVYNTIDEDQFSSRVISRGFLK 707
 Qy 678 ETKNSTRQTQVLGHTQTC--TLKLQPLNCIEDPVSPIVLRNPSL--VGTPLSAFNL 733
 Db 708 ENNERCLOKTMIVSQAQRCSVIIHIQEPS---DIISPLNLCWNTSLENPQT----- 756
 Qy 734 RVLAEADAQRALTALFPFENKCNNDNICODLSITF-----SFMSLDCLVVGPRFNVTV 789
 Db 757 NPALAEYSATKVSFIPFHKDCGDDGVCISDLVNLVQQLPATQQQPFIVSNQNKRLTFSV 816
 Qy 790 TVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASST-EVSGALKSTS 848
 Db 817 QLNKKEAYNTEIIVDFSENLF-----ASMSMPVDGTEVTCIASSQKSVT 864
 Qy 849 CSINHPFPENSEVETNITFDVDSKASLGNKLLKANVTSENMPRTNTEFOLELPVKY 908
 Db 865 CNVGVPAKSKQCVTTFINFDNLQ-NLQNAQISPRALSQENMADNSVNLKSLLY 923
 Qy 909 ANYVMVTHGVSTKYNLNTFASNTSRVMOHQVSNLQOR-----SLPISLVFLV 958
 Db 924 DAEIHIT-RSTNINIFYEVLSDGNVSVV-HSFE-DIGPKFTFSIKVTGVSVPVNSA--- 976
 Qy 959 PVRLNQTIVDRPQVTFSEN---LSSTCHTK-----RLPSHSDFLAR- 998
 Db 977 -----SVIHIPOYTKDNPLMYLTGVHTDQAGDISCAEINPLKIGTSSSVFSPKSEN 1030
 Qy 999 LRKAPVYVNSIAVCORIQCDDIPFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTAEI- 1057
 Db 1031 FRHIKELNCKRTASCNIMCWLRLQKGEVFLNVSTRIMWGTFAASTFOTVQLTAAAEID 1090
 Qy 1058 LFNDSVFTL-----LPGQAFVRSQTEKVEPFE-VPNPLPLVGVSSVGGLLLLALITA 1110
 Db 1091 TYNPQIVYIEENTVITP-----LTIMPKHEKVEVPTGVIGSVIAGILLLLALVA 1140
 Qy 1111 ALYKLGFFKROVKDM 1125
 Db 1141 ILWKLGFFKRYKXDM 1155

RESULT 12

IT2A_MOUSE STANDARD; PRT; 1178 AA.
 AC Q62469; Q62163;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Mus.
 OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
 RX MEDLINE=94363406; PubMed=8081889;
 RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
 RA Danjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
 RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
 RT not virus binding.";
 RL Cell Adhes. Commun. 2:131-143(1994).
 RN (2)
 RP SEQUENCE OF 450-1178 FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94355691; PubMed=7521231;
 RA Wu J.E., Santoro S.A.;
 RT "Complex patterns of expression suggest extensive roles for the alpha
 RT 2 beta 1 integrin in murine development.";
 RL Dev. Dyn. 199:292-314(1994).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
 CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
 CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
 CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
 CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
 CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VFMA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 229987; CA82877.1; -
 CC DR EMBL; X75427; CA853178.1; -
 CC DR PIR; S44142; S44142.
 CC DR HSP; P17301; IAOX.
 CC DR MGI; 96600; Itga2.
 CC DR InterPro; IPR000413; Integrin_alpha.
 CC DR InterPro; IPR002035; VWF_A.
 CC DR Pfam; PF01839; FG-GAP; 3.
 CC DR Pfam; PF00357; Integrin_A; 1.
 CC DR Pfam; PF00092; vwa; 1.
 CC DR PRINTS; PR01185; INTEGRINA.
 CC DR PRINTS; PR00453; VWFADOMAIN.
 CC DR SMART; SM00191; Int_alpha; 4.
 CC DR SMART; SM00327; VWA; 1.
 CC DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC DR PROSITE; PS0234; VWA; 1.
 CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Platelet; Signal; Repeat; Calcium; Magnesium.
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 1178
 CC FT DOMAIN 27 1129
 CC FT TRANSMEM 1130 1151
 CC FT DOMAIN 1152 1178
 CC FT REPEAT 42 100
 CC FT REPEAT 185 375
 CC FT DOMAIN 431 483
 CC FT REPEAT 485 546
 CC FT REPEAT 548 607
 CC FT REPEAT 612 664
 CC FT CA_BIND 496 504
 CC FT POTENTIAL.

FT CA_BIND 560 568 POTENTIAL.
 FT CA_BIND 624 632 POTENTIAL.
 FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1154 1158 GPRK MOTIF.
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 677 734 BY SIMILARITY.
 FT DISULFID 786 792 BY SIMILARITY.
 FT DISULFID 862 873 BY SIMILARITY.
 FT DISULFID 1016 1047 BY SIMILARITY.
 FT DISULFID 1052 1057 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;
 Query Match 18.2%; Score 1072; DB 1; Length 1178;
 Best Local Similarity 28.1%; Pred. No. 2.5e-64;
 Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;
 Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRYVVGAPQEIIVAANQSGSLYQC--DYST 54
 Db 27 YNVLGPKAIFSGPSSEQFGYSVQQLTNPQGNMLLVGSPWSPFPENRMGDVYKCPVDLPT 86
 Qy 55 GSEPIRLQ-----VPVAVNMSLGLSLAATSPPOLLAGCTVHTQTSENTYVYKGLC 107
 Db 87 ATCEKLNLSASISNVTETKTNMSLGLTLTRNPGTCGFLTCGPLWAHQCGNQYATGIC 146
 Qy 108 FLFGSLNRQQPO---KEPEALRGCPQSDIAFLIDGSGSIIHDPRMKEFVSTWMEQL 164
 Db 147 ----SDVSPDFQILTSFPAVQCPSL--VDVVVCDSESNYIP--WEAVKNEFLVKFTGL 199
 Qy 165 K--KSKTLFSLQMYSEFRIFHTFKFQNNPNSRLVKPITQLLIG-RTHATGVRKVKIRE 221
 Db 200 DIGPKKTOVALIOYANEPRIIFNLNDFETKEDVMVQATSETRHGGLDNTNFRALFEPARDY 259
 Qy 222 LLNITNGARKNAFKILVIDGEKFGDPLGYEDVIPADREGVIRYVIGV-----GDAFR 276
 Db 260 AYSQTSGRGRGATKVMVWVVDGESH-DGSKLKTVIQCNDBDEILRFGLVGLNRLNLD 318
 Qy 277 SEKSRQELNTIASPPRDHVFQVNNFEALKTIONLREKIFAIEGTQTGSSSFEHMSQ 336
 Db 319 TKNLKEIKAIATPTTERTYFFNVADAALLEKAGTLGEQIFSGIEGTVOG-GDNPQMEMAQ 377
 Qy 337 EGFSA--AITSNGPLLSVGSYDAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
 Db 378 VGFSAADYAPQNDILMLGAVGAFDWSGLTVQETSHKPVIFPKQAFDQVLQDRNHSSFLGYS 437
 Qy 392 AAILNRNOSVLGAPRYQHIGLVAFRONTGWMESNANV---KQTGAYFGASLCS 447
 Db 438 VAAISTEDGVHVFAGAPRANVTGQIVLYSVNK---QGNVTVIQSHRGDQGSYFGSVLCS 494
 Qy 448 VDSDNSGTDVLVIGAPHYEQTR--GGQVSVCPPLPRGQARWQCDVAVLYGEOQPGRF 505
 Db 495 VDVDKOTITDVLVGLGADTYNDLKEEGKYLFTITKGLNLQHQ---FLEGPEGTGNARF 551
 Qy 506 GAALTVLGVNGDKLTDVAIGAPEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPR- 564
 Db 552 GSAIAALSDINMGDFNDVI VGSFVENENSGAVIYNCHQGT-ITKYSQKILSGNGAFRR 610
 Qy 565 -LQYFGSLGSGQDLTWDGLVLDLTVGAGHVLRLRSQPLVLRVKAIMEFNPVARNVFE 623
 Db 611 HLQPFGRSLDGYGLDNGDSITDVSIGALGVQIQLWSQSIADVAIEALFTP----- 660
 Qy 624 NDQVWKAGEVRVCLHVQKSTRDLRREGIQSVVTVYDIALD-----SGRPHSAVENET 679
 Db 661 -DKITLLNKAOKITLKLCFRAEPRPQGNQV--AILFNMTLADGHSRVRTSRGVREN 717

Qy 680 KNSTRQTVLGLTQTCET--LKQLPNCIEDPVSPILVRLNLSFSLVGTPLSAFENLRPVL 737
 Db 718 SERFLOKMMVNEVQCKSEHHISIQKPS---DVVNPLDLAVDLSLENPGTS-----PAL 768
 Qy 738 AEDAQRLFTALPFPEKNCNDNICQDDLSI-----TFSPMSLDCLVVGPREPNVT 788
 Db 769 EAYSETVKVFSIPFYKECGSDGICISDLILDVQOLPAIQTSF-----IVSNQNKRLTFS 823
 Qy 789 VTVNRNDEGSYRTQVTFFFPDLDSYRKVSTLQNRQSRWRLACESASST-EVSGALKST 847
 Db 824 VILKNGESAYNTVLAEFSENLF-----ASFMPVDGTEVTCEVSSQSKSV 871
 Qy 848 SCINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTE--FQLELP 905
 Db 872 TCDGVGPALKSEQQVTFITNFDNLQ-NLQNAQAINFAQPSSEQ--ETNKADNSVSLTIP 928
 Qy 906 VKYAVVYVTVSHGVSTKYLNTASENTSRVMQHQYQVSNLQOR-----SLPISLV 955
 Db 929 LLYDAELHLT-RSTNINFIYSIDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVMA 984
 Qy 956 FLV-----PVRNLTQVWDRPQVTF--SNLS 980
 Db 985 LVTIHIPOYTKENKPNLLYLTGIQTDQAGDISCTAEINPLKPHTA----PSVFKNENFR 1040
 Qy 981 STCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIPFGIOBEFNATLKNLSFDW 1040
 Db 1041 ---HTKE-----LDCRTTSCNITCWLKDLHMAEYFINVTRVWNR 1080
 Qy 1041 YIKTSHNLLIVSTAILFNDSVFTLLPGOGAFVRSQETETKVEFFEPVNPPLIVGSSVG 1100
 Db 1081 FAASTQTQVLTAAASIDITHNPQLFVIEENAVTIPLIMKPKTEKAEVPT--GVIIGSIIA 1138
 Qy 1101 GLLLALITAAALKYLGFFKQYKDM 1125
 Db 1139 GILLLLAMTAGLWKLGFRRQYKDM 1163
 RESULT 13
 ID ITA2 HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 DE ITGA2.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
 RC TISSUE=Endothelial cells;
 RX MEDLINE=89308879; PubMed=2545729;
 RA Takada Y., Hemler M.E.;
 RA "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
 RT (platelet GP1a): homology to other integrins and the presence of a
 RT possible collagen-binding domain.";
 RT J. Cell Biol. 109:397-407(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP Rieder M.J., Amel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
 RA Rajkumar N., Toth E.J., Yi O., Nickerson D.A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RX X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
 RX MEDLINE=98019223; PubMed=9353312;
 RA Emelley J., King S.L., Bergelson J.M., Liddington R.C.;
 RT "Crystal structure of the I domain from integrin alpha2beta1.";
 RL J. Biol. Chem. 272:28512-28517(1997).
 RN [4]
 RP VARIANT HPA-5 (BR).

RX MEDLINE=94043762; PubMed=7901236;
RA Santoso S., Kalb R., Walka M., Kiefel V., Mueller-Eckhardt C.,
RT Newman P.O.;
RT "The human platelet alloantigens Br(a) and Br(b) are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2).";
RL J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10744142;
RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;
RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction.";
RL Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NAT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1; -;
DR EMBL; AF512556; AM34795.1; -;
DR PIR; A33998; A33998
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MIM; 192974; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; P:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
KW SIGNAL 30 1181
FT CHAIN 30 1181 INTEGRIN ALPHA-2.
FT DOMAIN 30 1181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1133 1154 POTENTIAL.
FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1155 1161 INTERACTION WITH HPS5.
FT REPEAT 45 103 FG-GAP 1.
FT REPEAT 7 2 FG-GAP 2.
FT REPEAT 188 378 VWFA.
FT REPEAT 378 433 FG-GAP 3.
FT REPEAT 434 486 FG-GAP 4.
FT REPEAT 488 549 FG-GAP 5.
FT REPEAT 551 610 FG-GAP 6.
FT REPEAT 615 667 FG-GAP 7.
FT CA_BIND 499 507 POTENTIAL.
FT CA_BIND 563 571 POTENTIAL.
FT CA_BIND 627 635 POTENTIAL.
FT SITE 1157 1161 GPFKR MOTIF.
FT DISULFID 83 92 BY SIMILARITY.
FT DISULFID 680 737 BY SIMILARITY.
FT DISULFID 789 795 BY SIMILARITY.
FT DISULFID 865 876 BY SIMILARITY.
FT DISULFID 1019 1050 BY SIMILARITY.
FT DISULFID 1055 1060 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 534 K -> E (IN ALLOANTIGEN HPA-5B;
dbSNP:1801106).
FT TURN 170 171 /FTid=VAR_003977.
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT STRAND 337 340
FT HELIX 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
Query Match 18.2%; Score 1069; DB 1; Length 1181;
Best Local Similarity 27.0%; Pred. NO. 4e-64;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

GO: GO:0007160; P: cell-matrix adhesion; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; WFA_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; WVFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; WFA_1.
 DR PROSITE; PS00442; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS0234; WFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22
 FT CHAIN 23 1167
 FT DOMAIN 23 1122 INTEGRIN ALPHA-10.
 FT TRANSMEM 1123 1145 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1146 1167 POTENTIAL.
 FT REPEAT 38 97 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 7 2 FG-GAP 1.
 FT REPEAT 2 2 FG-GAP 2.
 FT DOMAIN 167 350 WFA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LEU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;

Query Match 18.1%; Score 1065.5; DB 1; Length 1167;
 Best Local Similarity 29.3%; Pred. No. 6.8e-64;
 Matches 358; Conservative 196; Mismatches 488; Indels 181; Gaps 46;

QY 1 ENLDTENAMTQENARG-FQGSVVQLQGRS---VVVGAQEIIVAAQNRSLYQC----- 50
 DB 23 FNLDEHPLRFPQPPAEAGYSVLQVHVGQQRWMLVGAQWDGSGDRRGDVTRCPVGGAH 82
 QY 51 -----DYSTG-SCSPIRLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCS 99
 DB 83 NAPCAGHLGDYQLGNSHP-----AVNHLGMSLLETDCDGGFMACAPLWSRACGS 134
 QY 100 NTYVKGCLFLFGLNLRQOKFPQALRGCPQEDSDIAFLDGSGLIIPHD-----PRM- 153
 DB 135 SVFSGICARVDASFQPGQSLAPTAQR-CPTY-MDVVIVLDGNSNIYPMWSEVQTFLLRL 192
 QY 154 -KEFVSTVMEOLKSKTSLSLMOYSEFRHFTFEFQNNPRSLVKPITOLLGR-THT 211
 DB 193 GKLFIDP-EQIQ-----VGLVOYGESPVHWSLGDGFRTKKEVVRRAKNLSRREGRETK 245
 QY 212 ATGVKRVIRELLNITNGARKNAFKILIVITDQKGF-GDPLGYEDVPEADREGVIRYIG 270
 DB 246 AQAIMVACTEGFSQSHGRPEARLLVVVTDGSHDGEELPAALKACEAGR--VTRYGIA 303

RESULT 15

ITAL RAT

ID ITAL RAT

AC F18614;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

QY 271 V-GDAFRSEKS-----ROELNTIASKPPRDRHVQVNNFEALKTIQNLREKIFAIEGTQTG 325
 DB 304 VLGHYLRQRDPSSFLREIRTIASDPDERFFNFVNTDEAALTIDVDALGDRIFGLEGSHAE 363
 QY 326 SSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAAGGVFLYTSKSEKSTFINNTRVDS---- 381
 DB 364 NESSFGLEMSQIGFSTHRLKDGLIFGMVGAYDMGCSVLWLEGGHRLPPRMALEDEFPFA 423
 QY 382 -DMNDAYLYGA-AAIILNRVQSLVLCAPRYQHLGLVAMPR-ONTGMWESNANVKGTOIG 438
 DB 424 LQNHAAVLYGYSVSMLLRGRRLFLSGAPRFRHKGKVIAPOLKDKGAVRVAQSGQBOIG 483
 QY 439 AYFASCLSDVDVDSNGSTDLVLI GAPHY--EOTRGQGVSVCPRLPRGQARWQCDVLYG 496
 DB 484 SYFSELCPDTRDGTIDVLLVAAPMFLGPQNKETGRVYVYL--GQSLTLTQGTLOP 541
 QY 497 EQGQPMWRFGAALTIVLGVNGDKLTDVAIGAPEGEDNRGAVYLFHGTSGSGISFSHSORI 556
 DB 542 EPPQD-ARFGFAMGALPDNLQDGFADVAVGAPLEDHGQALYLYHGTQ-SGVRPHPAQRI 599
 QY 557 AGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAGHVLRLSOPVLRVKAIMEFNPREV 616
 DB 600 AAASMPHALSYFGRSVDRGLDLDGDDLDVAVAGAAIILSSRPVHLTSPSLEVTPOAI 659
 QY 617 ARNVFECNDQVVKGEAG--EVRVCLHVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRA 674
 DB 660 SVVQDCCR---RQGEAVCLTAALCFQVTSRTPGRWDH---QFYMRFTASLDEMTAGARA 713
 QY 675 VFNET--KNSTRQTVGLTQTCETLKLQIKNCIEDPVSPIVLRNLSVGTPLSARGN 732
 DB 714 AFDGSGQRLSPRLRLSVG-NVTCEQLHFHVLDTSDYLRPVALVTVTALDNTTKPG--- 768
 QY 733 LRPVLAEDAQRLFTALPFPEKNCNDNICQDLSITFSFMSLDC-----LVVGGPRE 784
 DB 769 --PVLNEGSPSTIQKLFPFKDCDPCNECVTDLVQ---VMDIRGSKAPFVVRGGRK 823
 QY 785 FNVTVVRNDCGDSYRTQVTFPFDLDSYRKVSTLQNRORSWRLACESASSTEVSGAL 844
 DB 824 VLVSTTLNRKENAYNTSLIIFSNL---HLASLTPQR-ESPIKVECAAPSA----- 872
 QY 845 KSTSCSINHPIFPENSEVTFNITFDVDSKASLG---NKL-----LKANVTSENMPRTN 896
 DB 873 HARLCSVGHVPFQTAGKVTFLLEFEFSCSLLSQVFGKLTASSDLSLRNGTLQENTAQT- 931
 QY 897 KTEFOLEPLVKYAVYVTVSHGVSTKYNLFTASENTSRVMQHQY-----Q 941
 DB 932 -----SAYIQYEPH-----LLFSESTLHRYEHPYGTLPVGPGEPEKPTTLR 973
 QY 942 VSNLG---QRLPISLVLFLVP-----VRLNQTIVDRPQVTFSENLSSTCHTKERL 989
 DB 974 VQNLGCVVVGIIISA--LLPAVAHGNYFLSLQVI-----TNWASCIQNLTEP 1022
 QY 990 PSHSDFLAELRKAPVNVNCIAVCQRIQCDIPFFGQEEFNATLKNLSFDWYIKTSHNHL 1049
 DB 1023 PGPPVHPBELQHTNRLNGSNTQCQVVRCHLGAKGTVEVSGLLRLVHNEFPRAKFKSL 1082
 QY 1050 LIVSTABELFNDSTVFLIPGQAGFVRSQTEKVEFEVNPPLPLIVGSSVGLLALLIT 1109
 DB 1083 TVVSTFELGTBEGSVLQLTEASRWSESLLEV-VQTRPILISLWILIGSVLGLLLALLV 1141
 QY 1110 AALYKLGFF-----KROYK 1123
 DB 1142 FCLWKLGFPAHKXIPEEKREK 1164

STANDARD;

PRT; 1180 AA.

[illegible]

Search completed: November 25, 2003, 14:17:38
Job time : 13.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMMSGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4387	74.7	1151	11 Q9J130	Q9J130 rattus norv
2	3907.5	66.5	1036	11 Q8CA73	Q8CA73 mus musculus
3	3815.5	64.9	920	6 Q28984	Q28984 sus scrofa
4	3484	59.3	1169	4 Q8IVA6	Q8IVA6 homo sapien
5	3326.5	56.6	1169	11 Q9QXH4	Q9QXH4 mus musculus
6	3251.5	55.3	1161	11 Q9QYE7	Q9QYE7 rattus norv
7	1529.5	26.0	1161	11 Q9MTV4	Q9MTV4 mus musculus
8	1519	25.9	1160	11 Q9R200	Q9R200 mus musculus
9	1414	24.1	1196	13 Q98TF1	Q98TF1 cyprinus ca
10	1363.5	23.2	1187	13 Q98TF0	Q98TF0 cyprinus ca
11	1350.5	23.0	1086	4 Q96HB1	Q96HB1 homo sapien
12	1270	21.6	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1163.5	19.8	1167	11 Q8B340	Q8B340 rattus norv
14	1119	19.0	1167	11 Q8B341	Q8B341 rattus norv
15	1051	17.9	1171	13 Q42094	Q42094 gallus gall
16	1041	17.7	1038	11 Q8BS01	Q8BS01 mus musculus

17	1014	17.3	895	11 Q9WUF8	Q9WUF8 mus sp. itg
18	1013.5	17.3	1160	6 Q8MKF4	Q8MKF4 felis silve
19	865	14.7	348	4 Q8TES5	Q8TES5 homo sapien
20	852	14.5	1332	5 Q8BPQ8	Q8BPQ8 halocynthia
21	807	13.7	205	11 Q83001	Q83001 rattus norv
22	753.5	12.8	780	13 Q06271	Q06271 xenopus lae
23	738	12.6	823	4 Q8WY18	Q8WY18 homo sapien
24	686.5	11.7	823	11 Q8CE84	Q8CE84 mus musculus
25	669	11.4	1032	11 Q61989	Q61989 mus musculus
26	643	10.9	1036	11 Q81YD5	Q81YD5 mus musculus
27	640.5	10.9	1033	6 Q9BGU3	Q9BGU3 bos taurus
28	630.5	10.7	257	11 Q8C270	Q8C270 mus musculus
29	619.5	10.5	1041	5 Q9UB90	Q9UB90 lytechinus
30	614.5	10.5	1041	5 Q76378	Q76378 lytechinus
31	579.5	9.9	1054	5 Q9U6S1	Q9U6S1 strongyloce
32	579	9.9	1034	13 Q98TT7	Q98TT7 gallus gall
33	550	9.4	1033	13 Q42598	Q42598 xenopus lae
34	534	9.1	1016	13 Q91779	Q91779 xenopus lae
35	530	9.0	974	11 Q924W2	Q924W2 rattus norv
36	529	9.0	1073	11 Q8CC06	Q8CC06 mus musculus
37	526	9.0	1047	6 Q9MZD6	Q9MZD6 bos taurus
38	525.5	8.9	1007	6 Q9GK48	Q9GK48 bos taurus
39	512.5	8.7	1034	6 Q9TUN4	Q9TUN4 oryctolagus
40	509.5	8.7	1036	6 Q9TUN6	Q9TUN6 sus scrofa
41	506.5	8.6	1012	11 Q70304	Q70304 mus musculus
42	491.5	8.4	1049	5 Q8SY51	Q8SY51 drosophila
43	475	8.1	833	5 Q9BPQ7	Q9BPQ7 halocynthia
44	469.5	8.0	1036	6 Q9TU44	Q9TU44 canis famil
45	468.5	8.0	604	11 Q8BQ25	Q8BQ25 mus musculus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
ID Q9J130
AC Q9J130
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrila K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5_
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02234; VWF_A; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.7%; Score 4387; DB 11; Length 1151;
Best Local Similarity 73.1%; Pred. No. 3e-313;
Matches 831; Conservative 146; Mismatches 158; Indels 2; Gaps 2;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLGSTRVVAAPQEVKAVNOTGALYQCDYSTNRCDEPI 60
17 FNLDTENAMTFQENARGFGQSVVQLGSTRVVAAPQEVKAVNOTGALYQCDYSTNRCDEPI 76
Db

```
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVYKGLCLFGLSGNLRQOQK 120
DB 77 PLQVPEAVNMSLGLSLAATTVPOLLACGPTVHQNCKENTYVNGLCYLFGLSGNLRKQOQ 136
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHDFRMKFEFVSTVMBQKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQESINAFLLDGGSGINTIDFKMKFEFVSTVMDQFQSKTLFSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRLNLTNGARKNAFKILIVI 240
DB 197 RTHFTFNDPKRNPDPKSHVRPQLNGRTKTASGIRKVVRELFQKINGARDNAKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQV 300
DB 257 TDGEKFGDPLNEDVPEABEAGIRYVIGVGNAPFKPQSRRELDTIASKPAGDVFQVD 316
QY 301 NFEALKTIONQLOEKIFAIEGTQTSSEFHEMSQEGFSAAITSNGLPLSTVSGYDWAG 360
DB 317 NFEALNTIRNLOEKIFAIEGTQTSSEFHEMSQEGFSAAITSNGLPLSGVSGFDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYPSKDKASFINTRIDSDMNDAYLGASAVISRNVRQSLVGLGAPRYQHIGLVAMFK 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRCGGQSVVCPL 480
DB 437 QNFGAWEPHTDIKSGOISYFGASLCSVDMDADGNTNLLIGAPHYETRCGGQSVVCPL 496
QY 481 PRGORARMOCDVAVLXGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWOCEAILHGDOQHPMGRFGAALTVLGDVNGDKLTDVAIGAPGEQENQAVYIF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLRLRSQ 600
DB 556 HGASVASISITPSHSQRIAGRFGLQYFGQSLSGGKDLTMDGLMDLAVGAQGRLLIRAQ 615
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660
DB 616 PVLRLKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 675
QY 661 YDLALDGRPHRAVNETKSTRRTQTVLGLTQTCETLKLQPLNCIEDPVSPIVLRNLF 720
DB 676 YDLALDGRSVVRAPFDETKNGILRIRVFGLTQKCTLKLPLDCVDNDSVSPILRLNY 735
QY 721 SLVGTPLSAFNLRPVLADAOQLFTALPFPEKNCNDONI CODDLSITFSPMSLCLVVG 780
DB 736 TLVGEPLRSSDLRPLVLAEMAQRIETAMPFPEKNCNDNTICODDLSITVSTRVNTLVG 795
QY 781 GPREFNVTVTRNDGSDSVRTQVTFPEPLDLVSRKVSTLQNSORSORSLACESASTEV 840
DB 796 DSRDFDVSLTRNDGSDSVGTVKTCYYPGSLVSRKVSASQNFSSKPKWRVIAE-PSSSEG 854
QY 841 SGALKSTCSINHIPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTKTEF 900
DB 855 QGVLKSTIWDINHPIIPANSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTKTEF 914
QY 901 QLELPVKYAVMYVTSKHGVSSTKYLNFTASENTSRVMHQYQVSNLQORSLPLSLFLPVPV 960
DB 915 QLELPVKYAIYVMTSGESSIKYLNFTASEMTSKVIHQYQVSNLQORSPFVSVFWFIPV 974
QY 961 RLNOTVMDRPOVTFSENLSSTCHTERLPSHSDFLAELKAPVNCSTAVCQRTQCDIP 1020
DB 975 QINKVTWIDPPQVTFSENLSSTCHTERLPSHSDFLAELKAPVNCSTAVCQRTQCDIP 1034
QY 1021 FFGIOBEFNATLKNLSFOWYIKTSHNLLIYSTABEILFNDSDSVFTLLPQCGAFVRSQTE 1080
DB 1035 SFNSKXEIENVTLQGNLLFDWYIETSHDHLVSTABEILFNDSDSAFALLPQETFVKAQTE 1094
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPGAEPO 1137
DB 1095 KVEPYTVHNPVPLIVGSSVGGVLLALITAGLYKLGFFKQYKDMMSGEGPGAEPO 1151
```

```
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
ID Q8CA73
AC Q8CA73;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; --
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;

Query Match 66.5%; Score 3907.5; DB 11; Length 1036;
Best Local Similarity 66.4%; Pred. No. 4.6e-278;
Matches 756; Conservative 121; Mismatches 142; Indels 119; Gaps 2;

QY 1 FNLDTENAMTPOENARFGQSVVQLQGRVVGAVPQEIIVAAANRGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTPOENARFGQSVVQLQGRVVGAVPQEIIVAAANRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVYKGLCLFGLSGNLRQOQK 120
DB 77 PLQVPEAVNMSLGLSLAATTVPOLLACGPTVHQNCKENTYVNGLCYLFGLSGNLRKQOQ 136
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHDFRMKFEFVSTVMBQKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQESINAFLLDGGSGINTIDFKMKFEFVSTVMDQFQSKTLFSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRLNLTNGARKNAFKILIVI 240
DB 197 RTHFTFNDPKRNPDPKSHVRPQLNGRTKTASGIRKVVRELFQKINGARDNAKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQV 300
DB 257 TDGEKFGDPLNEDVPEABEAGIRYVIGVGNAPFKPQSRRELDTIASKPAGDVFQVD 316
QY 301 NFEALKTIONQLOEKIFAIEGTQTSSEFHEMSQEGFSAAITSNGLPLSTVSGYDWAG 360
DB 317 NFEALNTIRNLOEKIFAIEGTQTSSEFHEMSQEGFSAAITSNGLPLSGVSGFDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRCGGQSVVCPL 480
DB 437 ENFGTWEPHTSIKG----- 450
QY 481 PRGORARMOCDVAVLXGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 451 ----- 450
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLRLRSQ 600
DB 451 -----SQRII GAHFSPLQYFGQSLSGGKDLTMDGLMDLAVGAQGRLLIRAQ 498
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660
DB 499 PVLRLKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 558
```



```
Db 259 TDGRKQGDNIYSYDVIPIMAEAAIIIRYAIAGVKAIFYNEHKSQELKAIASMPSEHYVFSVE 318
Qy 301 NFBAALKTIONQLEKIPAIETGTTGGSSSEHEHMSQEGFSAATISNGPLLSITVGSVDWAG 360
Db 319 NFDAKDIEQLEKIPAIETGTTGGSSSEHEHMSQEGFSAATISNGPLLSITVGSVDWAG 378
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAIAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 379 GAFLYPSNMPTFINMSQENEDMDAYLGYSTALAFWKGVHSLILGAPRHOHTGKVIFT 438
Qy 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQVSVCLP 480
Db 439 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDSGTDVLVIGAPHYEYQTRGGQVSVCPM 498
Qy 481 PRGORARWQCDVLYGSGQGWGRFGAALTVDGVNGDKLTDVAIGAPGEDNKGAVLYF 540
Db 499 P-GVGSRWHCCTTLHGSGQHPWGRFGAALTVDGVNGDSLADVAIGAPGEENRGAVIYF 557
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPORISASQIPRIQYFGQSLSGGQDLTMDGLVDLAVGSKGRVLLRTR 617
Qy 601 PVLRVKAIMFNPREVARNVFEQNDQVVKGEAGEVRVCLVHVKSTRDRRLREGIOISVVT 660
Db 618 PILRVSPVHTFAEISRSVFECQEQVAPEQTLSDATVCLHIESPKTQL--GDLRSTVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRQVQLGTCTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 FDUALDHGRILSTRAIFKETKTRALTRVKTGLGNKHCSVKLLUPACVEDSVTITLRLNF 735
Qy 721 SLVGTPLSAGFNLRPLVAEDAQRALTALFPFEKNCNDNIQQDLSITFSFMSLDCLVVG 780
Db 736 SLVGVPISSLQNLQPLAVDDQTYFTASLPFEKNCADHICQDDLSVVFDPDLKTLVVG 795
Qy 781 GPREFNVTVTRNDGDSYRQVTFPPFLDLVYKYSTLQ-----NQRQSRWR 829
Db 796 SDLELNDVTVSNDGDSYGTVTFLFYVGLSFRRAEAGVFLRKKBQDQWQRGQSHL 855
Qy 830 LACESASSTEVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLWSTGSSRHVIFRGGSQMTFLTFDVPSPKAEGLDRLLRARVGE 913
Qy 890 NNPRNTKNTFQLELPVKYAVYVMTVSHGVSTYKLNFTASE--NTSRVMQHYQVSNLQOR 948
Db 914 NNVPCTKNTFQLELPVKYAVYVMTVSHGVSTYKLNFTASE--NTSRVMQHYQVSNLQOR 973
Qy 949 SLPLSVLPVRLNQTVIWDROPVTFSENLSSTCHTKERLPKSHDPLAELRKAPVNVCS 1008
Db 974 DVPVSIWFMVPIELKGEAVN--TVMVSHPNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
Qy 1009 IAVCQRIQCDIPFGIOEEFNATLKNLSFDWIKTSHNHLILVSTAEIILFNDSVFTLLP 1068
Db 1033 IADCLHLRCDIPSLGIIDELFYFLKGNLSFGWISQTLQKVKVLLSEAEIIFNTSVYSQLP 1092
Qy 1069 GQGAFRVRSQETKVEPPEVENPLIIVGSSVGLLALLAIIAALYKLGKFFKROYKDMNSE 1128
Db 1093 GQBAFLRAQTKVLEMYKHNPNVPLIIVGSSVGLLALLAIIAALYKAGFFRQTKEMLEE 1152
```

RESULT 6

```
Q0QY67 PRELIMINARY; PRT; 1161 AA.
AC Q0QY67;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVlieten M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAP21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWF_A; 1.
DR Integrin.
KW
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.3%; Score 3251.5; DB 11; Length 1161;
Best Local Similarity 58.0%; Pred. No. 9.8e-230;
Matches 654; Conservative 158; Mismatches 302; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFQENARFGQSVQLOGSRVVGAPQEIIVAAVNRGSLVQCDYSTGSCPIR 61
Db 21 NLDVEPIVREDADAFSGQTVQFGSRLVVGAPLEAVNQTRGLYDCAPATGMQPIV 80
Qy 62 LQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVYKGLCFLFGSNLRQPPQKF 121
Db 81 LRSPLAEVNMSSLSLVATNNAQLLACGPTAQACVACVQNMVYAKGSCLLGSSSL-QFTQAV 139
Qy 122 PEALRCPCPEDSDIAFLIDGSGSIIIPDFRRMKKEFVSTWMEQLKSKTFLPSLMQYSEFR 181
Db 140 PASMPCEPRQEMDIAFLIDGSGSINQDFQAKMKDFYKALMGEPASTSTFLPSLMQYSNLK 199
Qy 182 IHPTFEFQNNPNRSLVVKPIITOLLGRTHRTATGVKVIKRELLNITNGARKNAFKILVIT 241
Db 200 THFTFTFEKNLDPQLSDVPIVQQLGTLTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259
Qy 242 DGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDHVFQVNN 301
Db 260 DGQYRDLPLEYSDVIPAADKAGIIRVAVIGDGAFAQEPALKELTNTIGSAPPQDHVFKVN 319
Qy 302 FEALKTIONQLEKIPAIETGTTGGSSSEHEHMSQEGFSAATISNGPLLSITVGSVDWAG 361
Db 320 FAALRSIQRLQEKIPAIETGTTGGSSSEHEHMSQEGFSAATISNGPLLSITVGSVDWAG 379
Qy 362 VFLYTSKEKSTFINMTRVDSMDNDAYLGAIAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 421
Db 380 AFLYPNTPTFINMSQENEDMDAYLGYSTALAFWKGVHSLILGAPRHOHTGKVIFTQ 439
Qy 422 NTCGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQVSVCLP 481
Db 440 EABHWPKESEVRGTQIGSYFGASLCSVDVDRDGSITDLVIGAPHYEYQTRGGQVSVFPV 499
Qy 482 RQGRARWQCDVLYGSGQGWGRFGAALTVDGVNGDKLTDVAIGAPGEDNKGAVLYF 541
Db 500 -GVRGRWQCEATLHGSGQHPWGRFGAALTVDGVNGDNLDVAIGAPGEESRGAVIYF 558
Qy 542 GTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 601
Db 559 GASRLIMPSPSORVTSQSLRLQYFGQSLSGGQDLTMDGLVDLAVGAGHVLRLSLP 618
Qy 602 VLRVKAIMEFNPREVARNVFEQNDQVVKGEAGEVRVCLVHVKSTRDRRLREGIOISVVT 661
Db 619 LLKVELSIRFAPMEVAKAVYQCHERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676
Qy 662 DLALDSGRPHSRVAFNETKSTRQVQLGTCTCETLKLQLPNCIEDPVSIVLRLNFS 721
Db 677 DLALDPRGLISRAIFDETKNCTLTGRKTLGLGDHCEVTKLLPDCVEDAVSPIILRLNFS 736
```

```
QY 722 LVGTPLSAFNGRLRPVLAEDAQLFTALPFPEKNCNDNI CDDLSITFSFMSLDCLVGG 781
DB 737 LVKDSASP-RNLHPVLAVGSDPHITASLPFEKNCCKOELLCEGLGISFNGLOVLYVGG 795
QY 782 PREFNVTVRNDGDSYRTQVTFPPDLDSVRKYSTLQNRORSQSWRLACESASSTVS 841
DB 796 SPFLVTVTVWNEGEDSVGLVKFYYPAGLSVRVTGQ-QPHQYPLRACEAPAAQED 854
QY 842 GALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENPRNTKTEFF 901
DB 855 --LRSSSCSINHPIFREGAKTTFMITFDVSKAFGLDRLLRKASSNNKPDNTKTAFA 912
QY 902 LELPKYAYVYVTVSHGVSTKYLNETASENTR-VMQHQYQVSNLQSRSLPISLVFLVPV 960
DB 913 LELPKYVTVTLISQEDSTNHNFSHGGRRQEAHRYRNNLSPLKLAVRVFWFVPV 972
QY 961 RLNTQVWDRPOVTFSENLS--TCHTKERLPSHSDFLAELRKAPVYVNGSIACQRIQCDI 1019
DB 973 LINGVAVMD--VTLSSPAQGVSCVQMKPPQNPDPFLTQIQRSVLDCSIADCLHFRCDI 1029
QY 1020 PFGIOEFPENATLKNLGFWDYIKTSHNHLIVTAEILFENDSVFTLLPGQAFVRSOTE 1079
DB 1030 PSLDIQDELDFLGNLSFGWYSQTLQEKVLLVSEAITFDTSVYSQLPGQEAFLRAQVE 1089
QY 1080 TKVEPEFVNPPLIVGSSVGLLALLALITAALYKLGFFKQYKDMW 1126
DB 1090 TTEEVYVYVEPIFLVAGSSVGLLALLITVVLVYKLGFFKQYKEML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; RAD25885.1; -.
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.0%; Score 1529.5; DB 11; Length 1161;
Best Local Similarity 33.8%; Pred. No. 4.3e-103;
Matches 396; Conservative 220; Mismatches 456; Indels 101; Gaps 37;
QY 1 FNLDTENAMTFQBNR-RFGQSVVQLQSGSRVVGAPQEIIVAAQNQRSLYQCDYSTGSCBP 59
DB 24 YNLDRTPQSLAQACRHFGYQVLIQIEDG-VVVGAPGE--GDNTGGLYHCRSTSEBFCOP 79
QY 60 IRLQVPEAVNNLSGLSLAATSPPOLLAGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQ 119
```

```
DB 80 VSLH-GSNHTSKVLGMLTATDAAGKSLACDPGLSRTCDQNTVLSGLCYLFPQSLGPM 138
QY 120 KFEALRGCEQSDSDIAFLIDGSGSIIPHDFRRKKEPVSTVMEQLKSKTKLFSMLQVSEE 179
DB 139 QNRPAYQECMKGVLDLVFLDFGSGSLDRKQFEKILEFMDKVMRKLSNTSQFAAVQFSTD 198
QY 180 FRIHFTPEF-QNNPNRSLVKPITOLLGRTHATGVKRIRELLNITNGARKNAFKILI 238
DB 199 CRTEFTLDVVKQKNPDVLLGSVQPMFLTNTFRAINVYVAHVKEESGARPDATKVLV 258
QY 239 VITDGEKFGDPLGYEDVPEADREG-----VIRYVIGVDADFRRSEKSOELNTIASKP 291
DB 259 IITDG-----EASDKGNISAAHDITRYIIGIGKHFVSQVKQKTLHIPASEP 304
QY 292 PRDHVFQNNPEALKTTQNLREKIPAI EGTQTGSSSSFEHMSQEGFSAAITNSGPLLS 351
DB 305 VEEFVKILTDFEKLKOLFQDQRRYIAIEGTRQDLTSFNMELSSSGISADLSKGHAVG 364
QY 352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGVAAA-IILNRNVQSLVLCAPR 409
DB 365 AVGAKWAGGFLLDREDLOGATFVGQPLTSDVVRGGLGTVAWMTSRSRPPLAAGAPR 424
QY 410 YQHIGLVAMPR--QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 467
DB 425 YQHVQVLLFQAPPEAGGRWQTKI EGTQIGSYFGGELCSVDLDQDGEAEALLIGAPLFF 484
QY 468 EQTRGGQVSVCLPRGORARWQDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGA 527
DB 485 GEQGRGVFTY--QRRQSLFEMVSELQGDGPGYPLGRFGAAITALTIDNGDRLTDVAVGA 541
QY 528 PGEEDNRGAVYLFHGTSGSGISPSHSORIASGKLSPRLOFGQSLSGGODLTMDGLVDLT 587
DB 542 PLEE--QGAVYIFNGKFG-GLSPQPSQRIQAGVFGIRWFGRSIHGKVDLGGDRLDADV 598
QY 588 VQAQGHVLLLRQPLRVKAIMFENPREARNVFCNDQVVKKEAG-EVRVCLRHVKST 646
DB 599 VGPEGRVVVLSRRPVDVVTLSFSPREIPVHEVECSVSAREEQKHGKVLKACFRKPLT 658
QY 647 RDRUREGOIQSVTYDIALDSGRPHSRVAFNETKNSTRQTVGLGTCTETLKLQLPNC 706
DB 659 PQ--FQGRLLANLTYTLQDHRMRSGLPFGDSHELSGNITSITP-DKSCLDHFHFHPIC 715
QY 707 IEDPVSPIVLRNFSLV--GTPLSAFGN-LRPVLAEDAQLFTALPFPEKNCNDNICQ 762
DB 716 IQDLISPIVSLNFSLEEETPRDQGRAMQPILRPSIHTV-TKEIPFEKNCGEDKKCE 774
QY 763 DDLSTITESFMSLDCLVVGGP-----REFNVTVTVRNDGEDSYRTQVTFPPDLDSYRKV 816
DB 775 ANLTLSPPARS-----GPLRLMSSASLAVETLNSNGEDAYVWRDLDPFRLSPRKV 827
QY 817 STLQNRQSRQSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTNITFDVDSKAS 875
DB 828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKLKNVSPSPFKAGQEVSLQVMTLLNSS 882
QY 876 LGNKLLLKANTYSEN-NMPTNKTEFQLEPLVKYAVYVTVSHGVSTKYLNETASENTR 934
DB 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNLITKEQENSTLYISFTPKGPKIQ 942
QY 935 VMQHOYQV---SNLQQRSLPISLVFLVPRLNQTVMIDRPO--VTFSENLS--TCHT 985
DB 943 QVQHVYQVRIQPSAYDNMP-TLEALVGVP-----WPHSEDPITYTWSVQTDPLVTC 995
QY 986 KE-RLPSSHDFLAELRKAPVYVNGSIACQRIQCDIPFGIQEEFNATLKNLSFDWYIKT 1044
DB 996 EDLKRPSSE--AEQPCLPGV-----QPRCPVIF---RREILIQVTGTVELSKELKA 1041
QY 1045 SHNHLIIYSTAEILLFENDSVFTLLPGQAFVRSOTETKVEFEVNPPLPLIVGSSVGG 1104
DB 1042 S-STLSLCSLSVSNSSKHFHLYGSKA-SEAQVLVVKVDLITHEKMLCHVYVLSGIGGLV 1099
QY 1105 LALITAALYKLGFFKQYKQYKDMW-SEGGPPGAEP 1136
```



```
Db 1100 LFLIFLALYKVGFFKRNLEKMEADGGVNGSP 1132
RESULT 8
Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065901; AAD25884.1;
DR HSP; P20701; ILFA.
DR MGB; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWA; 1.
DR PROSITE; PS02334; VWA; 1.
SQ SEQUENCE 1160 AA; 128127 MW; A33C531B139F1FAD CRC64;

Query Match 25.9%; Score 1519; DB 11; Length 1160;
Best Local Similarity 33.6%; Pred No. 2.5e-102;
Matches 394; Conservative 219; Mismatches 459; Indels 100; Gaps 36;

Qy 1 FNLDTENAMTFOENA-RFGGOSVVOLOGSRVVGAPQBIIVAAANQSGSYOCYSTGSCP 59
Db 24 YNLDRPTQSLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRSTSEFCQP 79
Qy 60 IRLQVPVEAVNMSGLSLAATTSPPQLACGPTVHOTCSNTYVYKGLCFGLSGNLRLQOPQ 119
Db 80 VSLH-GSNHTSKYLGMTTATDAAGKSLACDPGLSRTCDQNTYLSGLCYLPQSLGPM 138
Qy 120 KPFEALRGCPQEDSDIAFLIDGSGSII PHDFRRMKFEVTVMEQLKSKTKLFLSLMOYSEE 179
Db 139 QNRPAYQCMKGKVDLVEFDGSLDRKDFEKLIFEMKDVNRKLSNTSYQFAAVQFSTD 198
Qy 180 FRIHFTKFEF-QNNPNRSLVKPIITQLGRHTATGVRKVIARELNLNITNGARKNAFKILI 238
Db 199 CRTEFTFLDYKQKNPNPVLGSLVQPMFLNTFPAINVVAHVFKESGARPDATKVLV 258
Qy 239 VITDGEKFGDPLGYEDVTEADREG-----VIRYVGVGDAPRSEKSRQELNTIAKSP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIGIGHGFVSQVKQKTLHIFASEP 304
Qy 292 PRDHVFOVNNREALKTIQNLREKIFAIEGTQGTSSSPHEHMSQEGFSAITSNGLPLS 351
Db 305 VEEFVKIILDTFEKLKDLFTDQRRYYAIEGNRQDLTSFNMLSSSGISADLSKGHVVG 364
Qy 352 TVGSDWAGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRVQSLVGLGAPR 409
Db 365 AVGAKDWAGGDLREDLQATFVQCEPLTSDVRGCGYLYTVVAMTSSRSPLLAAGAPR 424
Qy 410 YCHIGLVAMFR--QNTGWMESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHY 467
Db 425 YQHVQVLLFQAPAEAGGRWNTQKIEGTQIGSYFGGLCSVDLDQDGAELLIGAPLFF 484
```

DR InterPro: IPR002035; VWF A.
DR Pfam: PF01839; FG-GAP; 5.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA_1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 24.1%; Score 1414; DB 13; Length 1196;
Best Local Similarity 31.3%; Pred. No. 1.4e-94;
Matches 373; Conservative 227; Mismatches 446; Indels 144; Gaps 39;

Qy 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQNRGSLYQCDYSTGS 56
Db 32 FNIDTEHPLRFNGAPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSGEMYSCTADLQS 88

Qy 57 CEPILRQVP-----VEAVNMSGLSLAATSPQLLACGPTVHQTSENTYVKGCLFLFGS 112
Db 89 CQ--RLRPGSESVRFPGMSAAVSSAALTS-----CGPYFPHECDGNSYLVNGVCYQFS 140

Qy 113 NLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFS 172
Db 141 SL-QAVSNFTAYQCSKREVNLFVFDGSSMKTVDFEMKNKFIKIMKLSNSIKFA 199

Qy 173 LMQYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELN-INGARK 231
Db 200 AVQFSDVTRVDFDNFYQSGSAEELMKE-THMKSILTNTHKAIDVILKNLNSMLSGADS 258

Qy 232 NAFKILIVTDGKFGDPLGYED--VIPADREGVIRYVIGDADFSEKSRQELNTIAS 289
Db 259 KAKALAVIITD---GDPSDNDYVNLKCDQNILRYIIGV-----KVDIETLQAS 309

Qy 290 KPRPDHVOVNFPEALKTQONLREKI PAIEGTQTGSSSSFEHMSQEGFSAAITSNGPL 349
Db 310 EPRNRTFYIKYSGLGKGLDNLQKIYNIESVDAQGRDQKELSQSGFSVYQESVI 369

Qy 350 LSTVGSYDWAGGVLYTSKESTFFINMTRVDSMN-DAYLGYAAAIILNRVQSVLVLGAP 408
Db 370 VGSVGSNDWRGALYEVTG-SGSDFKETETIDPAVNKDSYMGYSTVLGMRGVSLLFSGAP 428

Qy 409 RVQHIGLVAMFQNTQWGESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVE 468
Db 429 RAETGLVTLFTKNQNTWTVTNNGEQIGSYFGASLSLDDSDGSDFLVLGAPLFYQ 488

Qy 469 QTRGGQVVCPLPRGORARWQCDVAVLYGEO-----QPMGRFGAALTVLGDVNGDKLTDV 523
Db 489 SQ-----PRTEGLRYVYSLSEQYFQKTLNVQSSTTGRAAASVASLKDNLGDLSDV 540

Qy 524 AIGAPGEEDNRGAVYLFHGTSGSGISPSHS-QRIAGSKLSPLRLQVFGQSLSGQDLTMDG 582
Db 541 AVGAPLE--NEGYYVYVLGDRTHGINPELTQORISVQSVLPLQVQGVSLTQGMNDNDN 598

Qy 583 LVDLTVGAQHVLLLRQSVLRVKAIMEPNPREVARNVFECDNVVYKGEAGEVRVCLHV 642
Db 599 LTDIVIGAOGGIYLLKARPVMSYSAQLSPKSEISLNYFECPGS--NAPNANLTSCTFTV 656

Qy 643 QKSTRDLREGQIQSV--VTYDLALDGRPHSRVAFNETKNSTR--RQTVGLTLTCTET 698
Db 657 TERTSS---TGSLEKLNLSNLNVDRVGRMSRGFFDQSVSRITLQOQSVLLDSSGSCFN 713

Qy 699 LKQLPNCIEDPVSPILVRLNPLSVGTPLSAFGLNRPVLAEDAQRLFTALFPFKNCNGND 758
Db 714 ESIFMLRCVADTVSPVKIRWFS--QTEMLS-GNSVAVLDVHSRTEENVEVPQRCNSN 770

Qy 759 NICQDDLSTFFSFMISLDCLVWGPREFNVTVTVNRDGEDSYRTQVTFPDLDSYRKVST 818
Db 771 NSCVADLKLNFSTN-NTLVVENQAHTVQVSLANPCDDSYNTSIVLHYPEGISLSKFDA 829

Qy 819 LQNRQORSWRLACESASSTEVSGALKSTSCSINHPIFFENSEVTFNITFDV---DSKAS 875

Db 830 IKPSRTR-----SSCGDRDSCATNRTTCSIDLPPYRSGTTTQFLGTRFVWKWDNDSS 881
Qy 876 LGNKLLLKANVTSENNMPTNKTEFOLELPVKAVVAVVVTSHGV-STKYLNFTASENSTR 934
Db 882 NRMBEIMITANSNNNM---SDTEVRRSVVPQFAVDLAISLVAEDSVTYMNFSLSDRGPK 938
Qy 935 VMQHQYQVSNLQGRSLPISLVFLVPVR-----LN 963
Db 939 PLNITYKVENSGRKLDPVSVTLTLPCTPHVLLTPHTFSMHEVHHSFISSVHQIIMCLLN 998
Qy 964 QTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQ-----C 1017
Db 999 KHLFFSPSLSAVQVRSQSLREVEC-SQF--DLNKSSAVHFNLTADARLQNVKEYES 1055

Qy 1018 DIPFGIOEE--FNATLKGNLSPDWYIKTSHNHLIIVSTAEILLFNDSVFTLLPGQGAFAVR 1075
Db 1056 KYSEYERKONVFSISAEIINTNTSLYNQTS-----SELKYNPH-----R 1094

Qy 1076 SQTETKVEPPEVNPPLIV-GSSVGGLLLLALITAAALYKLGFFKQRYKD 1124
Db 1095 SQTETKVE-FVVPFSLMLIVCTGAVGGFFLLIILLLKCGFFKRNRPD 1143

RESULT 10
Q98TF0 PRELIMINARY; PRT; 1187 AA.

AC Q98TF0
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE C11-2.
GN CIA2.

OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB048537; BAB39135.1; --
DR HSSP: P20701; 1LFA
DR InterPro: IPR001969; Aspartatease_site.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 5.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 23.2%; Score 1363.5; DB 13; Length 1187;
Best Local Similarity 31.5%; Pred. No. 7.2e-91;
Matches 375; Conservative 209; Mismatches 451; Indels 157; Gaps 43;

Qy 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQNRGSLYQCDYSTGS 56
Db 32 FNIDTEHPLRFNGAPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSGEMYSCTADLQS 88

Qy 57 CEPILRQVP-----VEAVNMSGLSLAATSPQLLACGPTVHQTSENTYVKGCLFLFGS 112
Db 89 CQ--RLRPGSESVRFPGMSAAVSSAALTS-----CGPYFAHEDGNSYLVNGVCYQFS 140

Qy 113 NLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFS 172

Db 141 SL-QAVSNFTAAVOECSEKREVNLFVDFDSSMKAVFDMNKNFIKDWKMLSNSSIKFA 199
Qy 173 LMOYSEFRTHFTKEFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLN-ITNGARK 231
Db 200 AVQSTIRTVDFNDIQNSAEKLMKE-RHKSLTNTYKAINVLDKVNLSVSSGADP 258
Qy 232 NAFKILVITDGEKFGDPLGYED--VIPEADREGVIRYVIGVDGDAFSEKSRQELNTIAS 289
Db 259 NAQKALVIITD---GDPSSNDNDYILNICEQNILAYIIGV---KVDLTTLTQALAA 309
Qy 290 KPRDRHVQVNNFALKTIONLREKIFATEGTGTGSSSSFEHEMSQEGFSAITSNGPL 349
Db 310 EPKLNNFTYIQEYNGGLGLDNLQKTYNIEGSEKARQKELSQSGFVVYQESVI 369
Qy 350 LSTVGSVDMAGGVFLYT---SKEKSTFINTRVDSMDNDAYLVGAIAIILNRNVQSILVLG 406
Db 370 VGSVGSNDWKGALYEVNGSSKFKQTEITPAVN---KDSYMGSTVILGMRHGVSLFLSG 426
Qy 407 APRYOHTGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHY 466
Db 427 APRAEHTGLVTLFTKNSTVTVMRNINGEQIGSYFGASLSLLDSDSDSDFLLVGAPLF 486
Qy 467 YE-OTRG-QOVSVCPFLPRGRARQWQDAVLYGEO--QOPWGRFGAALTVLGVNNGDKLTD 522
Db 487 YQSOPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDLNGDGLSD 535
Qy 523 VATGAPCEEDNRGAVYLFHGTSGSGISPSHS-QRIAGSKLSPRLOYEGSISGGQDLTMD 581
Db 536 VAVGAPLE--NEGYYVITYLGDGTHGINPEHAPQIPARSVLGLQOQFVGSLSQGMNDND 593
Qy 582 GLVDLTVGAQCHVLLLRQPVLRKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLH 641
Db 594 NLPDIVIGTQGGIVLLNARPMVSVAQSLSNPHEISLNYFECPS--NAPNANFLTSCFT 651
Qy 642 VOKSTRDLREGQIQSV--VYTDALDSGRPHSAFNETKNSTR--ROTQVGLGLTQCE 697
Db 652 VTERTS--TGSLEKKLVNLSNLDVVRGMSRGFPDMPDSSSRTLQOQSVLLDSGSSCS 708
Qy 698 TLKQLPNCIEDPSPVILRNFS-----LVGTPLSARGNLRPVLAEDAQLFTALFFPEK 753
Db 709 NFSIFMLRCVADTVSPLKIRMFQSQMLSGNSLAVL-DIQSRTEEYEVVL-----FQR 761
Qy 754 NCGNDNICODDLSITFSFMSLDCLVGGPREFNVTVTRNDGDSYRTQVTFEFLDLSY 813
Db 762 NC-NSNSCVADLKLNFSTN-DTLVENVQAHFTVLVSLANPGDDSYNTSIVLHYPEGLSL 819
Qy 814 RKVSTLQORSORSRWLACESASSTEVSGALKSTCSINHPIPPENSEVTNITFDV--- 870
Db 820 SKFDAIKPSRTR-----SSCGDRDSGATNRTTCSINLPVYRSGLTTQFLGTFRTKW 871
Qy 871 DSKASLGNKULLKANVTSENNMPTNKTETOLELPVKYAVTMVVTSHGV-STKYLNPSTAS 929
Db 872 DYDWSDRMETITANSNDNGNM---SDMSVRRSIPQFAVELAISVAEDSVTYLNFSL 928
Qy 930 ENTSRVMOHQYVSNLQORSPLISLVLVPVRLNQVIMDRPQVTFSENLSSTCHTKERL 989
Db 929 DRGPKPLNIIYKVNVNLGLKLPVSVTLSPCQ-----THTVTLTPHNFMSQ 974
Qy 990 PSHSDFLAEALRKAPVANC-----SIAVCQRIQCDIPFFGQEE 1027
Db 975 EVHHSFISSVHQ--IIMCLLNKLHFFSPPELSAVQTRTIGRSLWC-----VSSISTGEI 1026
Qy 1028 FNATLKGNL-----SFDWYIKTSHNHLIVSTAELFNDVSNFTLLPG--OGA 1072
Db 1027 FRSSV--NLMAEAVLQNVKEYESKSYF-BFRDRHVENIS-AELNFTNSRYNQSTGLKYN 1082
Qy 1073 FVRSQTEKVEPFPVNPPLIVGVSSVGGLLALLIITAALYKLGFFPKQYKD 1124
Db 1083 PHRSQTEKVEFVIPPSPRMLIVCTGVAGVGFFFLIILILLKCGFFKRNRPD 1134

RESULT 11

Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DSC-2001 (T-EMBLrel. 19, Created)
DT 01-DSC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC008777; AA08777.1; -.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWA; 1.
DR Hypothetical protein.
KW SEQUENCE 1086 AA; 119223 MW; P6FF2546E8C632F9 CRC64;
SQ
Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.2%; Pred. No. 5.6e-90;
Matches 376; Conservative 186; Mismatches 432; Indels 175; Gaps 37;
Qy 1 FNLDENAMTFQ--ENARFGQSVQLOQSRVVGVVQAPQEIIVAAQNGSLYQCQDSTGSC 58
Db 26 YNLVDVGRASFSPPRAGRHFGRVQLV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81
Qy 59 PIRLOVPVEAVNMSLGLSLAATTPPQLLACAGPTVHTCSENTYVVKGLCFGLFGSNLRQP 118
Db 82 PVTLR-GSNVTSKYLQMTLA--TDP----- 103
Qy 119 QKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSE 178
Db 104 -----TDGS-----ILFAAVQFST 117
Qy 179 EPRHITFEKQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILI 238
Db 118 SYKTEPFDSDYVVKRKPDPALLKHVHMLLTNTFGAINYVATEVFRFELGARPDATKVL 177
Qy 239 VIITDGEKFGDPLGYDEIPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVQ 298
Db 178 IITDGE--ATDSNIDAAKD-----IIRYIIGIKGHFQTKESQETLHKFASKPASEFVKI 230
Qy 299 VNNFEALKTIONLREKIFATEGTGTGSSSSFEHEMSQEGFSAITSNGPLLSLVGSYDW 358
Db 231 LDTFEKLDLFTLEQKIVYIEGTSKQDLTSFNMELSSSGISADLSRGHAVGAVGAKDW 290
Qy 359 AGGVF-LYTSSEKSTFINTRVDSMDNDAYLVGAIAA--IILNRNVQSLVLGAPRYQHIGLV 416
Db 291 AGGFLLDKADQLDDTFIGNELTPEVRAGLYGVTVLWPSRQKTSLLASCAPRYQHMGVRV 350
Qy 417 AMFR--QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQ 474
Db 351 LLFQEPQGGGHWVQVQTIHGTQIGSYFGGELGVDDQDGETELLIGAPLFYGEQRGR 410
Qy 475 VSVCPPLPRGRARWQCDAV--LVGEGOPMGRFGAALTVLGDVNGDKLTDVAIGAPCEED 532
Db 411 VFYI-----QRRQLGFEEVSELOQDGYPLGRFGEAITALTDINGDLVDVAGAPLEE 464
Qy 533 NRGAVYLFHGTSGSGISPSHSQRISAGSKLSPRLOYEQSLSGGQDLTMDGLVLTGVAQG 592
Db 465 -QGVAVYIFNGRHG-GLSPQSPQRIEGTQVLSIQWFGRSIHGVKDLGDLGDLADVAVGAES 522

593 HVLLRSOPVLRVKAIMEFNPREVARNVFECNDQVW-KGKEAGBVRVCLHVQKSTRDLRL 651
 523 QMIVLSRPPVDMVTLMSFSPAEPVHEVECSYSTSNKMEGVNITICFOI-KSLIQP- 580
 652 EGQIQSVVTVYDALDSGRPHSRVFNETHKSTRQTVLGTQTCETLKLQLPNCIEDPV 711
 581 QGRLVANLTYTLQDGHRTRRGLFPGRHRLRNIAVT-TSMSCDTDFSFHPPVCVQDLI 639
 712 SPIVLRNFSL--VGTPLSAFGN-----LRPVLADAQRLFTALPFPEKNCNDNICQ 762
 640 SPINVSNFSLWEEGTPDRQAKDIPPIILRSLHSETWEI-----PFKEKCGEDKKCE 694
 763 DDLSTTFMSLDCLVVGCPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNO 822
 695 ANLRVSFSPARSRALRLTAFASLSVELSLNLEEDAYVQDLHFPFGLSPKRVEML--- 751
 823 RSORSWRLACGS--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 880
 752 KPHSQIPVSCPELPEESRLSRAL---SCNVSSPIFKAGHVALQMMFNTLVNSWGDSDV 808
 881 LKKNVTSENN---MPTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFETASENTSRVM 936
 809 ELHANVTNNEDSLLEDNSATTI---IPILYPINILIQDQEDSTLYVSTPKGPKIHQV 865
 937 QHOYQV---SNLQORSLP-ISLFLVLPVRLNQTIVDRPQVTFSENLSSTCHTK--ERLP 990
 866 KHMVQVRLOPSIHDHNIPTLEAVGVQPPSEGPITHQMSVQMEPPV--PCHYEDLERLP 923
 991 SHSD--FLAELRKAPVNCISAVCQRIQCDIPFGIOEFPNATLKNLSFDWYIKTSHNH 1048
 924 DAAEPCLPALFRCPVW-----FROEILVQVIGTLELVEGEIAS--SM 964
 1049 LLIYSTABILNDSVFTLLPGOGAFVRVQTEKVEPFPVNPPLIIVGSSVGGLLLALI 1108
 965 FSLCSLSISFNSSKHFHLYGSNASL-AQVMKVDVVYEQMLYLYVLSGGIGGLLLLL 1023
 1109 TAALYKLGFFKRYQKDMXSEG-GPPGAEP 1136
 1024 FIVLYKVGFRRNLKEMHAGRGVENGIP 1052

RESULT 12

Q8HZV0 PRELIMINARY; PRT; 927 AA.
 AC Q8HZV0:
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DE Lymphocyte function-associated antigen 1 (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovioidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Thumbikat P., Kannan M.S., Maheswaran S.K.;
 RT "Sequence of the alpha subunit of bovine lymphocyte function-
 associated antigen 1."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF40778; AAN63636.1; --
 FT NON_TER 1
 FT NON_TER 927
 SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match

Best Local Similarity 21.6%; Score 1270; DB 6; Length 927;
 Matches 339; Conservative 167; Mismatches 386; Indels 86; Gaps 28;

174 MOYSEFRHFTFKSQNNPRLSVKPTQLLQTHRTATGTVKRVIRELLNTNGARKVA 233
 4 VQFSTYFRTEFTLDYIQKDPDLAGLVKHMRLTLNTFGAINVYAKEVFRPDLGARPD 63

234 FKILVITOGKFGDPLGVEDVPIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIAKPPR 293
 64 TKVLIITDQK---PPTNTTLMPKPTSS-----RSLIGKGNFKTKESQEAHLQFASKPVE 116
 294 DHVQVNNFEALXTIQNLREKIPAFEGTQGTSSSSFEHMSQEGFSAATISNGPLSTV 353
 117 EFKYLDLDTFEKLKDLFTTELQKJYVIEGTSKQDLSFNMELSSSGISADISEGHGVVGA 176
 354 GSYDWAGGVF-LVTSKEKSTFINNTRVDSMDNDAYLGYAAA-IILNRVQSLVLGAPRYQ 411
 177 GAKDWAGGLDLKADLKSSFTFVGNELTVESRAGLYGTVTRLPSTRGTMSSLLATGAPKYQ 236
 412 HIGLVAMFRO--NTGMWESNANVKGQIGAYFGASLCSVDVDNSNGSTDLVLIGAPHYEQ 469
 237 HVGRVLLFQPKRGGPMWQIQEIDGQIGSYFGGELCGVDVDRDGETELLIIAAPLYYGE 296
 470 TRGOVSVCPLPRGARWOCDAVLGEOGQWGRFGAALTVLGDVNGDKLTVAIKAPG 529
 297 QRGGRVFIY---OKIQLEFQMVSELOGETGYLGRFGAAIAALTIDINGDELTDAVAGPL 353
 530 EEDNRGAVYLFHGTSSSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTVG 589
 354 EE--QGAVYIFNGQOG--GLSPRPSQRIEGTQMGFSGIOMFGRSIHGVKDLGGDLAVAG 410
 590 AQGHVLLRSOPVLRVKAIMEFNPREVARNVFECNDQVVKKEAG-EVRVCLHVQK--ST 646
 411 AEGQVIVLSSRPVVDIITSVSFSPAETPVHEVECSYSTSNKKEGVNLTVCFOVKSIIST 470
 647 RDLREQIQSVTVYDALDSGRPHSRVFNETHKSTRQTVLGTQTCETLKLQLPNC 706
 471 ----FOGHLVANLTYTLQDGHRTSRGLFPGGKHKLIGMTAVTPV-KSCFVFWHFHPIC 525
 707 IEDVPSIVLRLNFSL--VGTPLS--AFGNRLRVLAEADAQRLFTALPFPEKNCNDNIC 761
 526 IQDLISPINVLSLSLWEEGTPDRPALORDIPPILKPSHLETKEIPFKEKCGEDKNC 585
 762 QDLSITFSFMSLDCLVVGCPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQ 821
 586 EADLKLAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYVWQVTLSPFQGLSFRKVEIL-- 643
 822 QRSORSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 881
 644 -KPHSHVPVGCPELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKSGWDFIE 701
 882 LKANVTS-----ENMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFETASENTS 933
 702 LOANVSCNNESSLLEDNSATTI-----IPVMYPINVLTKQENSTLYISFTPKSPRI 754
 934 RVMQHOYQV---SNLQORSLPISLFLVLPVRLNQTVI---WD---RPQVTFSS-ENLSST 982
 755 HHVKHIYQVRIQPSNYDNMP-PLBALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESP 813
 983 CHTKERLPSSHDFLAELRKAPVNCISAVCQRIQCDIPFGIOEFPNATLKNLSFDWYI 1042
 814 SDEAE-----SCSFGT--EFCPCIDF--ROEILVQVMGMVELRGTI 850
 1043 KTSNNHLLIVSTABILNDSVFTLLPGOGAFVRVQTEKVEPFPVNPPLIIVGSSVGG 1102
 851 KAS-SMLSLCSSLAISFNSSKHFHLYHGNASM-AQVMKVDLVYVEKEMLYLYVLSGIGL 908
 1103 LLLALITAAALYKLGFFK 1120
 909 LLLFLIFIALYKVGFFKR 926

RESULT 13

O88340 PRELIMINARY; PRT; 1167 AA.
 AC O88340:
 DT 01-NOV-1998 (TRENBLrel. 08, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

```
DE Integrin alpha E1. (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020045; AAC23662.1;
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR DR
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.8%; Score 1163.5; DB 11; Length 1167;
Best Local Similarity 29.1%; Pred. No. 3.6e-76;
Matches 359; Conservative 211; Mismatches 458; Indels 205; Gaps 43;

Qy 1 FNLDTEA--MTFOENARGQSVVQLQGRVVVVGAPQEIIVAAHQGS-----LYQCDY 52
Db 20 FNMVDVNAVVTALQPGAPVLSLLHLDPSN-----NQTCLLVARRSSNNTAALYRCAL 74
Qy 53 STGCEPIRLQVPVEAVNMSLGLSLAAT--SPQLLAC-GPTVHQTCSENTYVYKGLCFL 109
Db 75 SI-SPDEIAC-PVEHICMPKRGYQGVTLGNHNGVLVCIQVQARKPSLSNLSLTGACSL 132
Qy 110 FGSNLROQPKFPALRG-----C-----PQE 131
Db 133 LTPNLDLQAQAFSDLEGFLDPAHVDSGDYCRSKGSTGEEKSARRRTVEEBDEED 192
Qy 132 DSDIAFLDGSIGSIIPDFFRMKEFVSTVMEQL--KSKTLFSLMOYSEBFRIHFTPEKF 189
Db 193 GTEAIVLDGSGSIEPDSDFQAKNFISTMRNFYKCFECNFALVOYGAIVQTFEDLOES 252
Qy 190 QNNPNRSLVKPITQLLGRHTATGVRKVIRELNIITNGARKNAFKLIVITDGEKFGDP 249
Db 253 RDINASLAKQSIQVQKEVTKTASAMQHVLNIFIPSRGSRKKALKVMVLTGDIQFQDP 312
Qy 250 LGVEDVTPADREGVIRYVVGDAFSEKSRQBLNTIASKPRDRHVFQVNNFEALTKIQ 309
Db 313 LNLTTVINSRPMQGVVRFAGVDAFKNNNTYRELKLIASDPKEAHTFKVTNYSALDGLL 372
Qy 310 NOLREKIFAETGTGSSSFEHMSOEGFSAATISNGP--LLSTVSGYVDWAGGVFLY-TS 367
Db 373 SKLQORIVHHEGT---VGDALQYLAQTGFSAQILDKGVLLGTGAFNWSGGALLYSTQ 429
Qy 368 KEKSTFTINMT-RVDS-DMNDAYLGYAAAIIILNRNVQSLVLGAPRYOHIGLVAMFRONTGM 425
Db 430 NGRCCFLNQAKEDSRVTQVSYLGSVLVHLKHAHGISYVAGAPRHLKRGAVFELRKEDR- 488
Qy 426 WESNA---NVKGTQIGAFYFASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQVSCPLPR 482
Db 489 -EEDAFVRRIEGEQMGVYFSGVLCVFDIDMDGTTDFLLVAAPFYHIRGEGRVYVYQVE 547
Qy 483 GQARWQCDALVYEQGQPMGRFCAALTVLGDVNGDKLTQVAIGAP-----GEEDNRGA 536
Db 548 -QDASFSLAHTLSGHPGLTNSRFGFAVAAVGDINQDKFTDVAIGAPLEGFGAGDCAVGS 606
Qy 537 VYLPFGTSGSGISPSHSQRAGSKLSRPRLOYFGQSLSGGODLTMDGLVLTGVAQGHVLL 596
```

```
Db 607 VIYNHSG-GLYDSSQQRASSASGLHYFMSVSGGLDFNGDGLADITVGSRDASVV 665
Qy 597 LRSQPVLRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKS---TRDRLREG 653
Db 666 LRSRPVVDLTVSMTFTP-----DALPMVFIGKM--DNLCCFEVDSSVVAASEPGLREM 715
Qy 654 QIQSVVYDLDALDSGRPHSRVFNENYKNSRTRQTVLGLTQTC----- 696
Db 716 FLNFTVDV-----TKQRQLQCEDESSGQCLRWKNGSFLCEHFMLI 760
Qy 697 ETLLQLPNCIEDPVSPIVLNLSLVGTPLSAFNLN-----PVLAEADQRLFTALF--P 750
Db 761 STEEL-----CEDCFNSITIKVTYE-----FOTSGGRRDYPNPTL--DHYKEPSAIFQUP 809
Qy 751 FEKNCGNINICQDDLSITFSFMSLCLVVGPPREFNVTVTRNDGEDSYRTQVTFPPFLD 810
Db 810 YEKDCNKKVFCIAEIQLTN--ISQELVVGTVKEVTMNIISLTNSGEDSYTNMALNYPN 868
Qy 811 LSYRKVSTLQNRQSRQSWRLACASSTEVSGALKSTCSINHPIPPENSEVTFNITPDV 870
Db 869 LQFKKI-----QKVPSPDVQCDPKPV--ASVLVNMCKIGHPII-KRSSVNVSVTWOL 918
Qy 871 DSKASLGNKLLKCANVTSNNMPRTNKTFFOLELPVKYAVYVMTVSHGVSTKYLNTASE 930
Db 919 EESVFNRTADITVTISNSNEKSLARETR---SLQFRHAFIAVLRS--PSVMYMN--TSQ 971
Qy 931 NTSRVNQHVQVSNLQGRSLPISLFLVPVRLNQTIVMDRPQVTPSPENLSST-----CHT 985
Db 972 SPDSHKKEFFNVHGENLFGAVFOLQICVPIKLODP-----QIVRVKNLTQDHTECTQ 1025
Qy 986 KERLPSPHFLAELRKAPVNVNCSIAVCQRICQDIPFGIOSEFNATLKGNSLFDWYIKTS 1045
Db 1026 SQBPACGSDPVQVQHWKSHVVCAL-----TSNKENVTVAAEISVG 1065
Qy 1046 HNHLLIVSTA-----EILFNDVSFTLLPGQAFVRSQTETKVEPF-----EVPNPLPLIV 1095
Db 1066 HTKQLLRDSELPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEETRSLPLII 1119
Qy 1096 GSSVGGLLILALITALYKLGFFKROYKQMMSE 1128
Db 1120 GSSIGGLLVLVIIALFKCGFFKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020046; AAC23663.1;
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19JUN03.*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	19	AAW65090 Human Beta-integrin
2	5868	99.9	1153	21	AAW07360 Human CD11b protei
3	5868	99.9	1153	23	ABG61469 Human Beta2 integr
4	5868	99.9	1153	23	AAU80252 Human integrin 1 a
5	5868	99.9	1153	23	AAO14428 Integrin Mac-1 alp
6	5858	99.7	1153	11	AAO4136 Alpha subunit of M
7	3464	59.0	1163	11	AAO7120 p150.95 alpha subu
8	3450	58.7	1163	19	AAW65091 Human Beta-integrin
9	3450	58.7	1163	21	AAW07361 Human CD11c protei

10	3450	58.7	1163	23	ABG61470 Human Beta2 integr
11	3448	58.7	1163	24	ABU07406 Protein differenti
12	3417	58.2	1161	16	AAW78166 Human beta-2 integ
13	3417	58.2	1161	18	AAW23049 Human beta 2 integ
14	3417	58.2	1161	19	AAW72825 Human alpha-d. Ho
15	3417	58.2	1161	19	AAW65089 Human Beta-integrin
16	3417	58.2	1161	19	AAW57491 Human Beta2 integr
17	3417	58.2	1161	20	AAW73342 Human alphad prote
18	3417	58.2	1161	21	AAW07359 Human alpha d clon
19	3417	58.2	1161	23	ABG61468 Human Beta2 integr
20	3401.5	57.9	1161	18	AAW23064 Human beta 2 integ
21	3401.5	57.9	1161	19	AAW72837 Human alpha-d deri
22	3401.5	57.9	1161	19	AAW65106 Human Beta-integrin
23	3401.5	57.9	1161	20	AAW73343 Human alphad prote
24	3401.5	57.9	1161	21	AAW07376 Human alpha d prot
25	3401.5	57.9	1161	23	ABG61485 Human Beta2 integr
26	3241.5	55.2	1161	16	AAW78169 Rat alpha-d subuni
27	3239.5	55.1	1161	18	AAW23082 Rat beta 2 integrin
28	3239.5	55.1	1161	19	AAW72834 Rat alpha-d #1. R
29	3239.5	55.1	1161	21	AAW60004 Rat alpha d polype
30	3239.5	55.1	1161	21	AAW07374 Rat alpha d protei
31	3239.5	55.1	1161	23	ABG61483 Rat Beta2-integrin
32	3232.5	55.0	1161	19	AAW65104 Rat beta-integrin
33	3232.5	55.0	1161	20	AAW73345 Rat alphad protein
34	3228	54.9	1161	18	AAW23061 Mouse beta 2 integ
35	3228	54.9	1161	19	AAW72836 Mouse alpha-d #2.
36	3228	54.9	1161	19	AAW65103 Mouse beta-integri
37	3228	54.9	1161	19	AAW60003 Mouse alpha d poly
38	3228	54.9	1161	20	AAW73347 Mouse alphad prote
39	3228	54.9	1161	21	AAW07373 Mouse alpha d prot
40	3228	54.9	1161	23	ABG61482 Mouse Beta2 integr
41	3224	54.9	1161	16	AAW78168 Mouse alpha-d subu
42	3218.5	54.8	1151	18	AAW23059 Rat beta 2 integrin
43	3218.5	54.8	1151	19	AAW72834 Rat alpha-d #2. R
44	3218.5	54.8	1151	19	AAW65101 Rat beta-integrin
45	3218.5	54.8	1151	19	AAW60001 Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090

ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

DT 28-SEP-1998 (first entry)

DB Human Beta-integrin CD11b subunit protein.

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX Homo sapiens.

XX OS

XX US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-0485618.

XX PR 07-JUN-1995; 95US-0485618.

XX PR 23-DEC-1993; 93US-0173497.

XX PR 05-AUG-1994; 94US-0286889.

XX PR 21-DEC-1994; 94US-0362652.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van DER VIEREN M;

XX

DR WPI: 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-1 diabetes

XX

PS Example 5; Fig 1A-D; 106pp; English.

XX

CC This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX

SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 19; Length 1153;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60

DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120

DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136

QY 121 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKKEFVSTVMEQKKSKTLFSLMQYSEEF 180

DB 137 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKKEFVSTVMEQKKSKTLFSLMQYSEEF 196

QY 181 RHFTKFQFONNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240

DB 197 RHFTKFQFONNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRVVGADAFRSEKSRQELNTIASKPPRDHVFQVN 300

DB 257 TDGEKFGDPLGYEDVPEADREGVIRVVGADAFRSEKSRQELNTIASKPPRDHVFQVN 316

QY 301 NFALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLSVTGSDWAG 360

DB 317 NFALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLSVTGSDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420

DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480

DB 437 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 496

QY 481 PRGORARQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

DB 497 PRGORARQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600

DB 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLKRVKAIMBNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660

DB 617 PVLKRVKAIMBNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720

DB 677 YDLALDSGRPHSRAVFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 736

QY 721 SLVGTPLSAFGNLRPVLAEQAORLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780

DB 737 SLVGTPLSAFGNLRPVLAEQAORLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVVRNDEGSYRTQVTFPPFLDLISYKVSITLQNRORSRNLACESASSTEV 840

DB 797 GPREFNVTVVRNDEGSYRTQVTFPPFLDLISYKVSITLQNRORSRNLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900

DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916

QY 901 QLELPVKVAVMVVTSHGVSITKYLNTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 960

DB 917 QLELPVKVAVMVVTSHGVSITKYLNTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 976

QY 961 RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELKAPVNVNCISIAVCQRIQCIP 1020

DB 977 RLNQTVIWDROPVTFSENLSTCHTKERLPSHDSFLAELKAPVNVNCISIAVCQRIQCIP 1036

QY 1021 FFGIOEEFNATLKGNLSPDVIKTSNHLILIVSTAEILFNDSVFTLLPGQGAFVRSOTET 1080

DB 1037 FFGIOEEFNATLKGNLSPDVIKTSNHLILIVSTAEILFNDSVFTLLPGQGAFVRSOTET 1096

QY 1081 KVEPEVENPDLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGAEPO 1137

DB 1097 KVEPEVENPDLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGAEPO 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX

AC AAB07360;

XX

DT 17-JAN-2001 (first entry)

XX

DE Human CD11b protein sequence.

XX

KW Human; macrophage infiltration inhibition; alpha d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type 1 diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX

OS Homo sapiens.

XX

PN WO200029446-A1.

XX

PD 25-MAY-2000.

XX

PF 16-NOV-1999; 99WO-US27139.

XX

PR 16-NOV-1998; 98US-0193043.

PR 08-JUL-1999; 99US-0350259.

XX

XX (ICOS-) ICOS CORP.

XX

XX Gallatin MW, Van Der Vieren M;

PI

XX WPI; 2000-387751/33.

DR

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites -

XX

XX Example 5; Fig 1; 270pp; English.

PS

XX Integrins are a class of membrane-associated molecules that participate

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins. Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

XX SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 21; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVWOLQGSRRVVVGAPQEIIVANQSGSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFQENARGFGQSVWOLQGSRRVVVGAPQEIIVANQSGSLYQCDYSTGSCPEI	76
Qy	61	RLQVPVAVNMSLGLSLAATSPPOLACGPTVHQTSCNTYVKGCLFGLFGSNLRQOPQK	120
Db	77	RLQVPVAVNMSLGLSLAATSPPOLACGPTVHQTSCNTYVKGCLFGLFGSNLRQOPQK	136
Qy	121	FPEALRGCPQSDSDIAFLIDGSGIIPHDFRMKEFVSTWQELKSKTLFSLMQYSEEP	180
Db	137	FPEALRGCPQSDSDIAFLIDGSGIIPHDFRMKEFVSTWQELKSKTLFSLMQYSEEP	196
Qy	181	RIHFTFEFQNPNSRLVLPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	240
Db	197	RIHFTFEFQNPNSRLVLPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	256
Qy	241	TGSEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN	300
Db	257	TGSEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN	316
Qy	301	NFEALKTIONQLREKIFAIEGTQTGSSSPHEMSQEGFSAIITNGPLISTVGSYDWAG	360
Db	317	NFEALKTIONQLREKIFAIEGTQTGSSSPHEMSQEGFSAIITNGPLISTVGSYDWAG	376
Qy	361	GVFLYTSKEKSTFINMTRVSDMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINMTRVSDMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR	436
Qy	421	QNTGMWESNANVKGTQIGAFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCPPL	480
Db	437	QNTGMWESNANVKGTQIGAFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCPPL	496
Qy	481	PRGORARWQCDVLYGSGQGPWGRFGAALTVLGVNNGDKLTDVAIGAPGEEDNRGAYLYF	540
Db	497	PRGORARWQCDVLYGSGQGPWGRFGAALTVLGVNNGDKLTDVAIGAPGEEDNRGAYLYF	556
Qy	541	HGTSGSGISPSHSORIASGLSRLQVFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ	600
Db	557	HGTSGSGISPSHSORIASGLSRLQVFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ	616
Qy	601	PVLRVKALMEFPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	617	PVLRVKALMEFPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	676
Qy	661	YDLALDSGRPHSRVAFNETKNSRRQTQVGLGTQTCETLKLQLPNCIEDPVSPIVLRINF	720
Db	677	YDLALDSGRPHSRVAFNETKNSRRQTQVGLGTQTCETLKLQLPNCIEDPVSPIVLRINF	736
Qy	721	SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG	780

Db	737	SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG	796
Qy	781	GPREFNVTVVRNDGSDSYRTQVTFPEPLDLSYRKVSTLQNRQSRQSWRLACASSTEV	840
Db	797	GPREFNVTVVRNDGSDSYRTQVTFPEPLDLSYRKVSTLQNRQSRQSWRLACASSTEV	856
Qy	841	SGALKSTSCSINHPIPIPEENSEVTNITFDVDSKASLGNKLLKANKVNTSENNMPTNKTEP	900
Db	857	SGALKSTSCSINHPIPIPEENSEVTNITFDVDSKASLGNKLLKANKVNTSENNMPTNKTEP	916
Qy	901	QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMOHQVQVSNLQORSPISLVFLVPV	960
Db	917	QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMOHQVQVSNLQORSPISLVFLVPV	976
Qy	961	RLNOTWIMDRPQVTFSENLSSTCHTKERLPSPSHDSFLAELRKAPVNVCSIAVCORIOCDIP	1020
Db	977	RLNOTWIMDRPQVTFSENLSSTCHTKERLPSPSHDSFLAELRKAPVNVCSIAVCORIOCDIP	1036
Qy	1021	FFGIQBEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET	1080
Db	1037	FFGIQBEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET	1096
Qy	1081	KVEPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ	1137
Db	1097	KVEPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ	1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX AC ABG61469;
XX DT 27-AUG-2002 (first entry)
XX DE Human Beta2 integrin alphaCD11b subunit.
XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX PN WO200230980-A2.
XX PD 18-APR-2002.
XX PF 15-OCT-2001; 2001WO-US32059.
XX PR 13-OCT-2000; 2000US-0688307.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX DR WPI; 2002-463260/49.
XX PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX PS Example 5; Page 191-194; 270pp; English.
XX CC The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
 CC spinal cord injury victim. The method also involves the use of
 CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
 CC molecule, vascular cell adhesion molecule). The method is useful for
 CC promoting locomotor recovery, inhibiting locomotor damage, limiting
 CC locomotor impairment, or limiting autonomic and sensory dysfunction
 CC following spinal cord injury. In particular, the spinal cord injury
 CC comprises compression of the spinal cord. The antibodies are also useful
 CC for reducing inflammation at the site of a central nervous system injury.
 CC The specification also details the identification of Beta2 integrin
 CC alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2
 CC integrins are implicated in diseases such as LAD (leukocyte adhesion
 CC deficiency, inflammatory response, diabetes, multiple sclerosis,
 CC arthritis, graft atherosclerosis, inflammatory bowel disease,
 CC Crohn's disease, ulcerative colitis, immune complex alveolitis
 CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
 CC sequence included for comparison with the Beta2 integrin alpha subunit
 CC sequences.
 XX
 SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 23; Length 1153;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATSPPLLACGPTVHQTCSNTYVVKGLCLFLGSLNLRQPOK 120
 DB 77 RLQVPVEAVNMSLGLSLAATSPPLLACGPTVHQTCSNTYVVKGLCLFLGSLNLRQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKGFVSTVMEQLKSKTFLSLMOYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKGFVSTVMEQLKSKTFLSLMOYSEEF 196
 QY 181 RIHFTPFQNNPNRSLVKPTITQLGRTHATGVRKVIKRELLNTNGARKNAFKILIVI 240
 DB 197 RIHFTPFQNNPNRSLVKPTITQLGRTHATGVRKVIKRELLNTNGARKNAFKILIVI 256
 QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQV 300
 DB 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQV 316
 QY 301 NFEALKTIQNLREKIPALIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
 DB 317 NFEALKTIQNLREKIPALIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVGLGAPYQHIGLVAMFR 420
 DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVGLGAPYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKGQTIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480
 DB 437 QNTGMWESNANVKGQTIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496
 QY 481 PRGQARWOCDAVLVYGEQCPWGRFGAALTIVLDVNGDKLTVDVAICAPGEEDNRGAVYLF 540
 DB 497 PRGQARWOCDAVLVYGEQCPWGRFGAALTIVLDVNGDKLTVDVAICAPGEEDNRGAVYLF 556
 QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQGHVLLRSQ 616
 QY 601 PVLRVKAIMEFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLRGQIQSVVT 660
 DB 617 PVLRVKAIMEFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLRGQIQSVVT 676
 QY 661 YDLALDSGRPHRAVNETKNSTRQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
 DB 677 YDLALDSGRPHRAVNETKNSTRQTQVLGTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEADAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
 DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
 QY 901 QLELPVYAVYVMVYTSHGVSYKLYNFNTASENTSRVMQHQVQVSNLQGRSLPISLVFLVPV 960
 DB 917 QLELPVYAVYVMVYTSHGVSYKLYNFNTASENTSRVMQHQVQVSNLQGRSLPISLVFLVPV 976
 QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
 DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036
 QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGFVRSOTET 1080
 DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGFVRSOTET 1096
 QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGPFKQYKDMMBEGGPPGAEPQ 1137
 DB 1097 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGPFKQYKDMMBEGGPPGAEPQ 1153

RESULT 4
 AAU80252
 ID AAU80252 standard; Protein; 1153 AA.
 AC AAU80252;
 DT 15-JUL-2002 (first entry)
 XX Human integrin 1 alpha-M subunit protein.
 DE Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
 KW inflammatory disease; autoimmune disorder; Crohn's disease;
 KW human immunodeficiency virus; HIV; myocardial infarction;
 KW Sjorgen's syndrome; rheumatoid arthritis.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 499..500
 FT /note= "Encoded by GCG CAG AGG"
 XX
 XX WO200218583-A2.
 XX
 XX 07-MAR-2002.
 PD
 XX 31-AUG-2001; 2001WO-US27227.
 XX
 XX 01-SEP-2000; 2000US-229700P.
 PR
 XX (BLOO-) CENT BLOOD RES INC.
 PA
 XX Springer TA, Shimoaka M, Lu C;
 PI
 XX WPI: 2002-382964/41.
 DR N-PSDB; ABK50046.
 XX
 XX Modified integrin-I or integrin I-like domain polypeptide useful as an
 PT immunogen to produce antibodies specific to polypeptide, comprises a
 PT disulfide bond such that polypeptide is stabilized in a desired
 PT conformation
 XX
 PS Disclosure; Page 109-112; 112pp; English.
 XX
 XX This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
 CC domain is stabilised in a desired conformation. The polypeptide of
 CC the invention may have anti-inflammatory or immunosuppressive activities.
 CC The polypeptides of the invention have an open conformation and are
 CC useful as immunogens to produce antibodies that selectively bind to
 CC integrin I-domain; and for identifying a modulator of integrin activity,
 CC or of interaction of an integrin and a cognate ligand. The polypeptide
 CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
 CC is useful for treating or preventing an integrin mediated disorder which
 CC is an inflammatory or autoimmune disorder in a subject and for
 CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
 CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
 CC infarction, Sjogren's syndrome, rheumatoid arthritis, dermatitis.
 CC A therapeutic composition comprising the peptide of the invention is
 CC useful for treating an integrin mediated disorder in a subject. The
 CC polypeptides and/or active or antigenic fragments are useful as
 CC reagents for diagnosis of integrin-mediated disorders. The present
 CC sequence represents the human integrin-1 alpha-M protein subunit used to
 CC generate the mutant polypeptides of the invention.
 XX
 SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 23; Length 1153;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOPQK 120
 DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOPQK 136
 QY 121 FPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMQVSEFP 180
 DB 137 FPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMQVSEFP 196
 QY 181 RIHFTKFEQNNPNPSLVKPIITQLLGRTHATGVKRVIRELLNITNGARKNAFKLILVI 240
 DB 197 RIHFTKFEQNNPNPSLVKPIITQLLGRTHATGIRKVRLEFNITNGARKNAFKLILVI 256
 QY 241 TDGEKEGDPGLGYEDVIAPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHFQVN 300
 DB 257 TDGEKEGDPGLGYEDVIAPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHFQVN 316
 QY 301 NFEALTKIQNLRKFIATGTTGSSSSFEHEMSQEGFSAAITSGNPLLLSTVGSYDWAG 360
 DB 317 NFEALTKIQNLRKFIATGTTGSSSSFEHEMSQEGFSAAITSGNPLLLSTVGSYDWAG 376
 QY 361 GVPLYTSKESKSTPINNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHTGLVAMFR 420
 DB 377 GVPLYTSKESKSTPINNTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHTGLVAMFR 436
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTREGQVSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTREGQVSVCP 496
 QY 481 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVLYLF 540
 DB 497 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVLYLF 556
 QY 541 HGTSGSGISPHSHSQRITAGSKLSPRLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 600
 DB 557 HGTSGSGISPHSHSQRITAGSKLSPRLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 616
 QY 601 PVLVRKAIMEFNEPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLVRKAIMEFNEPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVGLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 720
 DB 677 YDLALDSGRPHSRAVFNETKNSRRQTQVGLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPLVLAEDAQRLETAALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPLVLAEDAQRLETAALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 916
 QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENSTRVWQYQVSNLQGRSLPISLVFLVPV 960
 DB 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENSTRVWQYQVSNLQGRSLPISLVFLVPV 976
 QY 961 RLNQTVIWDPRQOVTSSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
 DB 977 RLNQTVIWDPRQOVTSSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
 QY 1021 PFGIOEEFNATLKGNLSDFWYIKTSHNHLIVSTAEILLFNDSVFLLPQCGAFVRSQTEF 1080
 DB 1037 PFGIOEEFNATLKGNLSDFWYIKTSHNHLIVSTAEILLFNDSVFLLPQCGAFVRSQTEF 1096
 QY 1081 KVEPEFVNPPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137
 DB 1097 KVEPEFVNPPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1153

RESULT 5

AAO14428
 ID AAO14428 standard; protein; 1153 AA.
 XX AAO14428;
 XX 03-MAY-2002 (first entry)
 XX Integrin Mac-1 alpha subunit.
 XX
 KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
 KW open conformation; integrin related inflammatory disorder;
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;
 KW viral infection; cancer; gene therapy; vaccine;
 KW bioactive agent screening.

Unidentified.

WO200204521-A2.

17-JAN-2002.

09-JUL-2001; 2001WO-US21805.

07-JUL-2000; 2000US-216600P.

(CALY) CALIFORNIA INST OF TECHNOLOGY.

(BLOO-) CENT BLOOD RES.

Springer T;

WPI; 2002-148167/19.

New integrin I domain protein having alteration in at least 2
 noncontiguous regions and exits in an open conformation, useful for
 treating, preventing or suppressing inflammatory or immunological
 disorders

Example 1; Fig 1F; 90pp; English.

The invention comprises structurally biased variant integrin inserted (I)
 domain proteins, wherein the alterations to the protein occur in at least

FT Modified-site 1045..1047
FT /*label= putative N-glycosylation site
FT Modified-site 1051..1053
FT /*label= putative N-glycosylation site
FT Modified-site 1076..1078
FT /*label= putative N-glycosylation site
FT region 1..16
FT /*label= signal peptide
FT region 1106..1134
FT /*label= putative transmembrane region
XX
PN EP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-0115159.
XX
XX 23-AUG-1988; 88US-0235353.
PR 09-MAR-1989; 89US-0321239.
XX
PA (DAND) DANA FARBBER CANCER INST INC.
XX
XX Springer TA, Corbi A;
XX WPI; 1990-125938/17.
DR N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
PT inflammation and viral infections, and in diagnosis
PT
XX Disclosure; Page ?; ?pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC recognition of and migration to sites of inflammation. It also attaches
CC to cellular substrates as part of this function making it useful in
CC visualising endothelial tissue.
CC Mac-1 is a member of the Integrin Gene superfamily.
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX
SQ Sequence 1153 AA;

Query Match 99.7%; Score 5858; DB 11; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1132; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 76

Qy 61 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSENTYVVKGLCFLFGSNLRQOPQK 120
Db 77 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSENTYVVKGLCFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQKSKTFLSLMQYSEEP 196

Qy 181 RIHFTFEFQNNPRSLVXKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFEFQNNPRSLVXKPIQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVIPAEDREGVIRYVIGVDAPFRSEKSRQELNTIASKPPROHVFOVN 300
Db 257 TDGEKFGDPLGYEDVIPAEDREGVIRYVIGVDAPFRSEKSRQELNTIASKPPROHVFOVN 316

Qy 301 NFEALKTIONOLREKI PAIEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKI PAIEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFFINMTVDSMDNDAYLGYAAAIIILNRNQSVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTVDSMDNDAYLGYAAAIIILNRNQSVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 496

Qy 481 PRGORARWOCDAVLVYGEQGPWGRFGAALTVLVDVNGDKLTVDVAIGAPGEDNRGAYVLF 540
Db 497 PRGORARWOCDAVLVYGEQGPWGRFGAALTVLVDVNGDKLTVDVAIGAPGEDNRGAYVLF 556

Qy 541 HGTSGSGISPSHSORJAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSORJAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616

Qy 601 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGOIQSVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGOIQSVVT 676

Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 736

Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHOYOVSNLQORSLSISLVLVVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHOYOVSNLQORSLSISLVLVVPV 976

Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCDIP 1036

Qy 1021 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGQCAFVRSOTET 1080
Db 1037 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGQCAFVRSOTET 1096

Qy 1081 KVEPPEVPNPPLIVGSSVGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
Db 1097 KVEPPEVPNPPLIVGSSVGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 7
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX AC
XX 25-MAR-2003 (updated)
DT 05-FEB-1991 (first entry)
XX
XX p150.95 alpha subunit encoded by clone lambdaX47.
DE
XX p150.95 leucocyte adhesion receptor alpha-subunit;
KW hairy cell leukaemia; rhinovirus.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 61..63
FT /label-glycosylation site
FT Modified-site 89..91
FT /label-glycosylation site
FT Modified-site 385..387
FT /label-glycosylation site

FT Modified-site 392..394 /label=glycosylation site
 FT Modified-site 697..699 /label=glycosylation site
 FT Modified-site 735..737 /label=glycosylation site
 FT Modified-site 899..901 /label=glycosylation site
 FT Modified-site 904..906 /label=glycosylation site
 FT Modified-site 939..941 /label=glycosylation site
 FT Modified-site 1050..1052 /label=glycosylation site
 FT Domain 1108..1133 /label=transmembrane
 FT Region 1..19 /label=signal peptide
 FT Region 20..44 /label=N-terminus

XX WO9010646-A.
 PN 20-SEP-1990.
 PD 09-MAR-1990;
 XX 02-FEB-1989;
 PR (DAND) DANA FARBER CANCER INST INC.
 PA Corbi AA, Springer TA;
 PI WPI; 1990-304985/40.
 DR N-PSDB; AAQ06068.
 XX Treatment of viral esp. rhino-viral infection - by admin. of alpha
 PT sub-unit of p150.95 cell surface adhesion receptor, opt. together
 PT with a beta chain of CD-18 family.
 XX Disclosure; Fig 3; 59pp; English.

CC Clone lambda X47 was isolated from a cDNA library constructed from
 CC total RNA extracted from phorbol myristate acetate stimulated Hu-60
 CC myelomonocytic cells. The library was screened with oligonucleotide
 CC probes based on tryptic peptide fragments of p150.95. The sequence
 CC can be attached to appropriate control elements and expressed in
 CC prokaryotic and eukaryotic cells. The protein can be used to treat
 CC or prevent rhinoviral infection because it interacts with ICAM-1
 CC and inhibits cell-virus attachment. It can also be used as an
 CC anti-inflammatory agent.
 CC See also AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1163 AA;
 SQ Query Match 59.0%; Score 3464; DB 11; Length 1163;
 Best Local Similarity 61.1%; Pred. No 3e-279;
 Matches 690; Conservative 138; Mismatches 295; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCYSTGSCBPI 60
 DB 20 FNLDTEELTAFRVDSAGFSDSVVQVANSWVVGAPQKITAANQGTGLYQCYSTGACBPI 79
 QY 61 RLQVPVEANVMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 120
 DB 80 GLQVPPEANVMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLGSLNLRQPOK 137
 QY 121 FPEARLGCQPOEDSLAFLDGGSGIIPHDFRMRKFEVSTVMEQLKKSKTLFSLMOYSEF 180
 DB 138 LPVSRQECRQODIVFLDGGSGISRRNFATMNFVRAVISQFORPSTQFSLMOFSNKF 197

QY 181 RIHFTPEFQNNPNBSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
 DB 198 QTHFTPEFRNTSNPLSLASVHQLQGFYTAITAIQNVHRLPHASVGARDATKILIVI 257
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQBLNTIASKPPDRDHVQVN 300
 DB 258 TDGKKEGDSLKYKDVIPMAADAAGIIRYAGVGLAFQNRNSWKLNDIASKPSQEHIPKVE 317
 QY 301 NFEALTIQNLREKIPAIETGTGSSSSFEHEMSQEGFSAAITSGNPLLSVGSVDWAG 360
 DB 318 DFDALDKIQNLREKIPAIETGTGSSSSFEHEMSQEGFSAAITSGNPLLSVGSVDWAG 377
 QY 361 GVFLYTSKEKSTFINMTRVDSDMNDAYLVYAAAILIILNRVQSLVGLGAPRYQHILGLVAMFR 420
 DB 378 GAFLYPNWSPPTFINNSQENVDMRDSYLGYSTELALWKGVQSLVGLGAPRYQHILGLVAMFR 437
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVVCP 480
 DB 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEYQTRGGQSVVCP 497
 QY 481 PRGORARWQCDVLYGEOGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
 DB 498 PRGWR-RWMCDAVLYGEOGQPGWGRFGAALTVLGDVNGDKLTDVWIGAPGEENRGAYLYF 556
 QY 541 HGTSGSGISPSHSORIASGKSLPRELOYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGVLPSPISPSHSORIASGKSLPRELOYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMBPNPREVARNVFECDQVYVKGEGEVRVCLHVQKSTRDLREGIQSVVT 660
 DB 617 PVLWVGVSQPIPAEIPRSAFECREQVVSBOTLVQSNICLYDKRSKNLLGSRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 677 LDALDPGLSPRATFOETKNSRVRVLGLKAHCENFNLLPSCVEDSVTPITLRLNF 736
 QY 721 SLVCTPLSACGNLRPLAEDAORLFTALFPKCNKGNDCNDCODDLSITFSFMSDCLVVG 780
 DB 737 TLVGKPLLAFLNLRPLAEDAORLFTALFPKCNKGNDCNDCODDLSITFSFMSDCLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
 DB 797 SNLELNAEVMVWNGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 854
 QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 900
 DB 855 SQGTWSTSCRINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 914
 QY 901 QLELPVKYAVVMVTSKGVSTKYLNFSTAS-ENTSRVMQHOYQVNSLQORSUPISLVFLVP 959
 DB 915 QLELPVKYAVVMVTSKGVSTKYLNFSTAS-ENTSRVMQHOYQVNSLQORSUPISLVFLVP 974
 QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPFSDHDLAELRKAPVAVNCISAVCORIQCDI 1019
 DB 975 VELNQEAVMVDVEVSHPNPSLRCSQKIAPASDFLAHIQKNPVLDCSAGCLFRCDV 1034
 QY 1020 PFGIQOEFEFNATLKNLSFDWYIKTSNNHLLIVSTAEILFNDSVFTLLPGGAFVRSQTE 1079
 DB 1035 PSFSVQBELDFTLKNLSFGWVRQILQKVSWSVAEITFDTSVYSQLPQGERFMRAQTT 1094
 QY 1080 TKVEPPEVNPPLPLIVGSSVGLLLALLITAAALYKGLFFKQYKDMSE 1128
 DB 1095 TVLEKYKVNPTPLIVGSSVGLLLALLITAAALYKGLFFKQYKDMSE 1143

RESULT 8
 AAW65091
 ID AAW65091 standard; Protein; 1163 AA.
 XX AC AAW65091;
 XX DT 28-SEP-1998 (first entry)
 XX

OS Homo sapiens.
PN WO200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US271139.
XX
PR 16-NOV-1998; 98US-0193043.
PR 08-JUL-1999; 99US-0350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX WPI; 2000-387751/33.
DR
XX
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous
PT system injury sites -
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This
CC sequence was used in an alignment to identify a novel beta2 integrin
CC alpha subunit: alpha_d (AAA60014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked
CC to aberrant alpha_d function e.g. Type I diabetes, atherosclerosis,
CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
CC adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal
CC antibodies may be used in the inhibition of macrophage infiltration at
CC the site of a central nervous system injury. The monoclonal antibodies
CC can also be used to detect and diagnose Crohn's disease.
XX
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3450; DB 21; Length 1163;
Best Local Similarity 61.0%; Pred. No. 4.4e-278;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDTENMTFOENARFGOSVVOLOGSRVVGAPQEIIVAAQORGLYOCDYSTGSCPT 60
DB 20 FNLDTEELTAFRVDSAGFSDSVVQYANSWVVGAPQKIITAAQIGGLYOCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSPLLACGPTVHCEGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMQYSEEF 180
DB 138 LPVSRQECPRQEDIVFLIDGSGSISRNFEATMNFVRAVISQFQRPSTQFSLMQFSNKF 197
QY 181 RHFTKEFQNNPNPSLVKPIITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
DB 198 QTHFTFEPRRTNPLSLASVHLQLOGFTYTATAIQNVVHRLFHASVGAARDIAKILIVI 257
QY 241 TDEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNITIASPPRDHVFQVN 300
DB 258 TDGKKGSDLDYKDVPMDADAGIIRYAIUGVLAQFNRSWKELNDIASKPSQEHFKVE 317
QY 301 NFEALKTIONLREKIPAEICTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAQ 360
DB 318 DFDALDKIQNLREKIPAEIGTETISSSSELEMAQEGFSAVFTPDGPVLGAVGSEFTWSG 377
QY 361 GVFLYTSKESKSTFINMTVRDSDMNDAYLGVAAAILLRNVQSLVLGAPRYQHIGLVAMFR 420

DB 378 GAFLYPPNMSPTPINMSQENVMDRDSYLGYSYSELALWKGVSQSLVGLGAPRYQHICKAVIFI 437
QY 421 QNTGHWESNANVKTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSCPL 480
DB 438 QVSRQWRMAEVIQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEYQTRGGQVSCPL 497
QY 481 PRGORARWOCDAVLYGEOGQPWGRFGAALTVLGVDVNGDKLTDAIGAPEGEDNRGAVYLF 540
DB 498 PRGWR-RWDCDAVLYGEOGQPWGRFGAALTVLGVDVNGDKLTDAVIGAPGEENRGAAYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGVLGPSISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRYKALMEFNPREVARNFECDNDQVVGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLWVGSMQFIPAEIPRSAFECEQVQVSEQLVQSNICLYIDKSKNKLGSRLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRTOVLGLGTCTETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 LDALAPGRLSFRAIFQETKRSLSRVVLGKHAECNPNLLSCVEDSVIPIILRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNICQDDLSITFSFMSDCLVWG 780
DB 737 TLVCKPLLAFLNRLPMLAALAQRYFTASLPFEKNCADHICQDNLGIFSFPGLKSLVVG 796
QY 781 GPREFNVTVVRNGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 SNLELNAEVMVMNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSJLHLC--CSAPVG 854
QY 841 SCALKSTSCSNHPIFFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSNNIPRISKITIF 914
QY 901 QLELPKYAVVMVTVSHGVSTKYNLFTAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 959
DB 915 QLELPKYAVIVVSSHQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISAVCQRIQCDI 1019
DB 975 VELNQEAVMQVDEVSHPNQPSLCSSEKIAIPASDFLAHQNPVLDCSIAGCLFRCDV 1034
QY 1020 PFGTQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSOTE 1079
DB 1035 PSFSVQEEUDFTLKNLSFGWVRQILQKVSVSVVAEILFDTSVYSQLPQGEAFWRAQTI 1094
QY 1080 TKVEPFEVNPPLIVGSSVGLLLALITLALYKLGFFKQYKDMSE 1128
DB 1095 TVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKVGFFKQYKEMEE 1143
RESULT 10
ABG61470
ID ABG61470 standard; Protein; 1163 AA.
XX ABG61470;
AC
XX
DT 27-AUG-2002 (first entry)
XX
DE Human Beta2 integrin alphaCD11c subunit.
XX
XX Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
OS Homo sapiens.
XX

PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
DR WPI; 2003-058520/05.
XX
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
PS Claim 1; Page 225-228; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3448; DB 24; Length 1163;
Best Local Similarity 60.9%; Pred. No. 6.5e-278;
Matches 688; Conservative 136; Mismatches 299; Indels 6; Gaps 4;
QY 1 ENLDTENAMTFOENARGFGQSVVVOLOGSRVVVGAPOEIVAAQNRGLYOCYSTGSCPEI 60
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 20 FNLDTTELTAFRVDSAGFGDSVVVQVANSVVVGAPOKITAANQTGGLYOCYSTGACEPI 79
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVYKGLCLFLGSLNRQQPOK 120
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMVLTGLCLLGLPT--QLTOR 137
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 121 FPEALRGCOQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKKSKTLFSLMQVSEEF 180
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 138 LPVSRQECPRQEQDIIVFLIDGSGSISSRNFAFMNFRAVISOQFORPSTQFSLMQFSNKF 197
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 181 RIHFTKEFQNNPNPSLKYPTOLLGRTHGTATGVRKVIPELLNITNGARKNAFKILIVI 240
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
XX

DB 198 QTHLTPEEPRTSNPLSLASVHLOQGYTATATQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDEKFGDPLGYEDVPEADREGVIRVYVIGVDAFRSEKSEQELINTIASPPRDHVPQVN 300
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 258 TDGKKEGDTLDYKDVIPMDAAGIIRIYAGVLAQFQNRNSWKELDIASKPSOEHIFKVE 317
QY 301 NFEALKTIQOLREKIFAIEGTOTGSSSPHEHMSOEGFSAAITSGNPLLSITVGSYDWAG 360
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 318 DFDAKDIQOLREKIFPIEGTETSSSEFELEWAQEGFSAVFTPDGVLGAVGSPMTWSG 377
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNRVQSLVGLGAPRYOHIGLVAMPR 420
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVGLGAPRYOHTGKAVIFT 437
QY 421 QNTGHWESNANVKTQIGAYFGASLSDVDNSGSTDVLIGAPHYVEQTRGGQVSCPL 480
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 438 QVSRQRMKAEBVTGTQIGSYFGPSLSDVDSDGSTDVLIGPHYVEQTRGAQVSCPL 497
QY 481 PRGORARWQCDVLYGEGQOPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLFL 540
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 498 PRGWR-RWNCDAVLYGEGQHPWGRFGAALTVLGVDNGDKLTDVIGAPGEENRGAYLFL 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 557 HGVLPSPISPSHSQRIAGSQLSRQYFGQALSGQDLTQDGLVDLAVGARGQVLLLR 616
QY 601 PVLRYKAIMFENPREVARNVFECDQVYVKGKEAGEVVRVCLHVQKSTRDRLBEGIOQSVWT 660
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 617 PVLWVGSMQFIPAEIPRISAFECREQVQVSEQTLLVQSNICLIYDKRSKSLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRTOVLGLTCTETLKLQLPNCIEDPVSPTVLRNLF 720
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 677 LDLDLDFGRLSPRATFQETKNRSRVRVGLKKAHCENFNLLPSCVEDSVTPITLRNLF 736
QY 721 SLVGTPLSAGFNLPRVLAEDAQRLLFTALFPFKNCNGNDNICQDDLSITFSFMSDCLVVG 780
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 737 TLVCKPLLAFLPRLAADAQRYFTASLPFKNCGADHICQDNLGIFSFPGLKSLLVG 796
QY 781 GPRFNVTVVRNDGDSYRTQVTFPPDLDSYKRYSTLQNRQSRWSRLACESASSTEV 840
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 797 SNLELNAEVMVMVNDGDSYGTITTFSPAGLSYRYVAEGQKQQLRSLHLCTDSAPVG-- 854
QY 841 SGALKSTSCSINHPIFFPENSESVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTFF 900
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 855 SQGTWSTSCRINHILIFRGGAQITFLATFDVSPKAVLGDRLLLLTANVSENNTPRTSKTFF 914
QY 901 QLELPVYAVYVMVYVTSYKYLNFNTAS-ENTSRVMQHOYQVSNLQORSUPISLVFLVP 959
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 915 QLELPVYAVYVYVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDI 1019
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 975 VELNQEAVVMVDVSLPQNLSRCSSEKIAQIPASDFLAHIQKNPVLDCSIAGLCRFRCDV 1034
QY 1020 PFFGIQEBFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSVFTLLPGGCAFVRSQTE 1079
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 1035 PSFSVQEBLDTFLKGNLSFGVWRQILQKKVSVVSAEITFTDTSYSQLPGQEAFWAQT 1094
QY 1080 TKVPEFVEPNPLPIVSGSSVGLLLALITAAALYKGLFFKQYKQKDMSE 1128
DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 1095 TVLEKYKVHNPTPLIVGSSIGGLLLALITAVLYKVGFQKQYKEMMEE 1143
XX
RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.
 XX OS Homo sapiens.
 XX PN US9831029-A.
 XX PD 03-NOV-1998.
 XX PF 07-JUN-1995; 95US-0482293.
 XX PR 07-JUN-1995; 95US-0482293.
 XX PR 23-DEC-1993; 93US-0173497.
 XX PR 05-AUG-1994; 94US-0286889.
 XX PR 21-DEC-1994; 94US-032652.
 XX PA (ICOS-) ICOS CORP.
 XX PI Gallatin WM, Van Der Vieren M;
 XX DR WPI, 1998-609318/51.
 XX DR N-PSDB; AAV67281.
 XX PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 XX PT immunohistochemical analysis
 XX PS Example 5; Column 61-66; 106pp; English.
 XX CC The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotype antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2). Antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX SQ Sequence 1161 AA;
 Query Match 58.2%; Score 3417; DB 19; Length 1161;
 Best Local Similarity 59.8%; Pred. No. 2.5e-275;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTQENARGQSVVQLQGSRRVVVGAPOETVAANORGLYQCDYSTGSCPEI 60
 DB 17 FNLDVEEPTIQEDAGGFGQSVVQFGGRLVVGAPLEVVAAVAAOTGRLYDCAATGMCOPI 76
 QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGCLPFLGSLNRQDPQK 120
 DB 77 PLHIRPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-EIIQT 135
 QY 121 FPEALRGCPQSDIAFLIDGSGIIPHDPRRMKEFVSTWEOQLKSKTLPSLMQYSEEF 180
 DB 136 VPDATPECPHQMIDIVFLIDGSGIDQDNFQNMKGFOVQAVMGQPEGDTLTPALMQYNLL 195
 QY 181 RIHTEFEFQNNPNSRLVKPIITOLGRTHTATGVRKVIKRELLNITNGARKNAKILIVI 240
 DB 196 KIHTFTQFRYSPSQSLSDPVLQKGLTFTATGILTVTQLFHKGARKSAKILIVI 255
 QY 241 TDGKFGDPLGVEDYIPADREGVIRYVIGVDAFRSEKSRQELNTASKPRDHVFOVN 300
 DB 256 TDGQKYKDPLEYSVIPAERKAGIRYVIGVDAFRSEKSRQELNTASKPRDHVFOVN 315
 QY 301 NFEALKTQNLQREKIPAEIQTGSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGSFWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 376 GAFLYPPNMSPTFINMSQENVDRDSYLGYSTELAWKGVQNLVGLGAPRYQHTKAVIFT 435
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYHYEQTRGGQSVCP 480
 DB 436 QVSRQWRKKAETGTQIGSYFGASLCSVDVDSNGSTDLVILGAPHYHYEQTRGGQSVCP 495
 QY 481 PRGORARWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTVDVAIGAPGEDNRYLYLF 540
 DB 496 PRGORVQWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTVDVAIGAPGEDNRYLYLF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLRLSL 615
 QY 601 PVLRVKAIMFENPREVARNVFECDNDQVVKGEKAGEVRVCLHVQKSTDRDLREGIOISVVT 660
 DB 616 PVLKVGVMRFPSPVEKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIOSSVR 673
 QY 661 YDLALDSGRPHSRAVFNETKNSRRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 674 FDLALDPRUTSRAIFNETKNPTLTKRKTGLGLGHCHETLKLPLDPCDEVVSPILHLNF 733
 QY 721 SLVGTPLSAGFNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGSQLPTASLPFEKNCNDNICQDDLSITFSFMSLDCLVVG 793
 QY 781 GPRFENVTVVRNDGEDSYRTQVTFPPDLVSRYKYSTLQNRQSRQSWRLACESASSTEV 840
 DB 794 SSLELNVITVMNAGDSYGTVSLVYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
 QY 841 SGALKSTCSINHPFIPENSEVFNITFDVDSKASLGKLLKXKLLKXKLLKXKLLKXKLL 900
 DB 853 EG-LRSLRSCSVNHPFIFHEGNGTGFIVFDVSKATLGDRLMRASASNNKASSKATF 911
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNF-TASENTSRVMOHQVQVSNLQORSPLSLVFLVP 959
 DB 912 QLELPVKYAVYVMTSHGVSTKYLNF-TASENTSRVMOHQVQVSNLQORSPLSLVFLVP 971
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAEIRKAPVNVCSIAVCORICDI 1019
 DB 972 VLLNGVAVMDVWMEAPSQL--PCVSRKPPQHSDFLTQISRSPLMDCSIADCLQPRCDV 1029
 QY 1020 PFGIOEENATLKGSLFDVYIKTSHNHLIYSTAEILFNDVSTLPGGAFVRSQTE 1079
 DB 1030 PSFSVQEELOFTLKGSLFGVNRVRETLOKKVLVWSVAEITFDSVYSQLPQOEFMRAQME 1089
 QY 1080 TKVEPPEFVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKRYQKDMSE 1128
 DB 1090 MVLEDEEVYNAIPIMGSSVGGALLLALITATLYKLGFFKRYKEMLED 1138
 RESULT 15
 AAW65089
 ID AAW65089 standard; Protein; 1161 AA.
 XX AAW65089;
 AC AAW65089;
 XX 28-SEP-1998 (first entry)
 DT 28-SEP-1998 (first entry)
 XX Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT

Qy 61 RLQVPEAVNMSLGLSLAATTSPQOLACGPTVHOTCSNTYVVKGLCFGLGSLNRQOQOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPQOLACGPTVHOTCSNTYVVKGLCFGLGSLNRQOQOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMBQLKSKTFLSLMQYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMBQLKSKTFLSLMQYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 316
Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEOQPNRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGQARWQCDVLYGEOQPNRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLGGQDLTMDGLVDLTGVAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLGGQDLTMDGLVDLTGVAQGHVLLLRQ 616
Qy 601 PVLRYKAIEMEPREVARNVFCNDQVVKGEVRVCLHVOKSTRPRLRGOQSVVT 660
Db 617 PVLRYKAIEMEPREVARNVFCNDQVVKGEVRVCLHVOKSTRPRLRGOQSVVT 676
Qy 661 YDLALDSDGRPSRAVFNETKNSRROTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSDGRPSRAVFNETKNSRROTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNRPVLADQRLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLADQRLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSORSRMLACESASSTEV 840
Db 797 GPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSORSRMLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 916
Qy 901 QLELPVKYAVVMVTSYKYNFTASNTSRVNHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVTSYKYNFTASNTSRVNHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNOTVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1020
Db 977 RLNOTVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1036
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNLLIYSTAEILFNDVSVFTLLPQGAFAVRSQTE 1080
Db 1037 FFGIOBEFNATLKNLSFDWYIKTSHNLLIYSTAEILFNDVSVFTLLPQGAFAVRSQTE 1096
Qy 1081 KVEPPEVPNPLIIVGSSVGGLLLLALITAALYKLGFFKRYKQKDMWSEGGPPGABPQ 1137
Db 1097 KVEPPEVPNPLIIVGSSVGGLLLLALITAALYKLGFFKRYKQKDMWSEGGPPGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVVOLQSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVOLQSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPQOLACGPTVHOTCSNTYVVKGLCFGLGSLNRQOQOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPQOLACGPTVHOTCSNTYVVKGLCFGLGSLNRQOQOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMBQLKSKTFLSLMQYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMBQLKSKTFLSLMQYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 316
Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDWAG 376

QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILRNRRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILRNRRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676
QY 661 YDLALDSGRHSRAVFNKSTRQVGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRHSRAVFNKSTRQVGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVRNDEGDSYRTQVTFPPFLDLVYKVKSTLQNRQSRQSWRLACASSTEV 840
DB 797 GPREFNVTVRNDEGDSYRTQVTFPPFLDLVYKVKSTLQNRQSRQSWRLACASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSRSLPISLVLVVPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSRSLPISLVLVVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGQEBFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGFVRSQTET 1080
DB 1037 FFGQEBFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KVEPFEVNPPLTVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLTVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 60
DB 17 ENLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVEAVNMSLGLSLAATTSPQQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
DB 77 RLOVPVEAVNMSLGLSLAATTSPQQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEFVSTVMEOLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEFVSTVMEOLKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFSEKSRQBELNTIASKPPRDHVPQVN 300
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFSEKSRQBELNTIASKPPRDHVPQVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILRNRRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILRNRRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600

Db 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLONORSQSRWFLACASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLONORSQSRWFLACASSTEV 856
QY 841 SGALKSTCSINHPIEPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTET 900
Db 857 SGALKSTCSINHPIEPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTET 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVLVFPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVLVFPV 976
QY 961 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFQWYIKTSHNHLIIVSTABILFNDVFTLLPGGAPVRSOTET 1080
Db 1037 FFGIQEEFNATLKNLSFQWYIKTSHNHLIIVSTABILFNDVFTLLPGGAPVRSOTET 1096
QY 1081 KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGABEQ 1137
Db 1097 KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGABEQ 1153

RESULT 4

US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEX: 312-474-0448
; TELEFAX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHTCSENTYVKGCLFPLFGLNLRQQPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHTCSENTYVKGCLFPLFGLNLRQQPQK 136
QY 121 PPEALRGCPQEDSDTAFILIDSGSIIPHDFRMKEFVSTVMEQLKSKTKLFSLMQYSEEF 180
Db 137 PPEALRGCPQEDSDTAFILIDSGSIIPHDFRMKEFVSTVMEQLKSKTKLFSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVVKPIITQLLGHTHATGVRKVIKRELLNITNGARKNAFKILVVI 240
Db 197 RIHFTFKFQNNPNRSLVVKPIITQLLGHTHATGVRKVIKRELLNITNGARKNAFKILVVI 256
QY 241 TDGEXFGDPLGVYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEXFGDPLGVYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQV 316
QY 301 NPEALKTIONQLREKI FAIEGTQTGSSSFHEHENSEGFSAAITNSGPLLSTVGSYDWAQ 360
Db 317 NPEALKTIONQLREKI FAIEGTQTGSSSFHEHENSEGFSAAITNSGPLLSTVGSYDWAQ 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 496
QY 481 PRGQBARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDAITGAPGEEDNRGAVLYF 540
Db 497 PRGQBARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDAITGAPGEEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLONORSQSRWFLACASSTEV 840

797 GPRFNVTVVRNDGSDSVRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 916
901 QLELPVKYAYVMVTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 960
917 QLELPVKYAYVMVTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 976
961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
977 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036
1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTABILLFNDVSFTLLPGOGAFVRQOTET 1080
1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTABILLFNDVSFTLLPGOGAFVRQOTET 1096
1081 KVPEFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKRYQKDMVSEGGPPGAEPQ 1137
1097 KVPEFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKRYQKDMVSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAQOEIVAAANQORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAQOEIVAAANQORGLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPVTHQTCSENTYVKGICFLFGSNLRQOPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPVTHQTCSENTYVKGICFLFGSNLRQOPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTMEOQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTMEOQLKSKTFLSLMOYSEEF 196
Qy 181 RHFTFKEFQNNPNRSLVKPITOLLGRTHATGVRKVRRELLNITNGARKNAFKILVI 240
Db 197 RHFTFKEFQNNPNRSLVKPITOLLGRTHATGVRKVRRELLNITNGARKNAFKILVI 256
Qy 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLYAAAILRNVRQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLYAAAILRNVRQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYTRGGQVQVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYTRGGQVQVCP 496
Qy 481 PRQARARWOCDAVLYGEOQPMGRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYL 540
Db 497 PRQARARWOCDAVLYGEOQPMGRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYL 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLVRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRQGIQSVVT 660
Db 617 PVLVRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRQGIQSVVT 676
Qy 661 YDLALDSGRPHSRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFEKNCNDNI CODDLSTTFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFEKNCNDNI CODDLSTTFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 976
Qy 961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036

QY 1021 PFGIOEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1096
QY 1081 KVEPEVFNPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1137
DB 1097 KVEPEVFNPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marehall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,859
; REFERENCE/DOCKET NUMBER: 127866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 76
QY 61 RLOQVPEAVNMSLGLSLAATTSPQLLAGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPQLLAGPTVHQTCSNTYVKGCLFLPGSNLRQPOK 136
QY 121 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKEFVSTVMEQLKSKTFLSLMQYSEEF 180

RESULT 7

US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

DB 137 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTPEFQNNPRSLVKPIITQLGRTHATGVRKIVIRELLNITNGAKNAFKILIVI 240
DB 197 RIHFTPEFQNNPRSLVKPIITQLGRTHATGVRKIVIRELLNITNGAKNAFKILIVI 256
QY 241 TDEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQELNITIASPPRDHVFQVN 300
DB 257 TDEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQELNITIASPPRDHVFQVN 316
QY 301 NFEALKTIONQREKIPAIETGOTGSSSFEHEMSQEGPSAAITNSGNPLSTVGSYOWAG 360
DB 317 NFEALKTIONQREKIPAIETGOTGSSSFEHEMSQEGPSAAITNSGNPLSTVGSYOWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRDLREGOIQSVYT 660
DB 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRDLREGOIQSVYT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRRTQVQLGTCTETLKLQPNCTEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRRTQVQLGTCTETLKLQPNCTEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFQNLPRVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFQNLPRVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHQYOVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHQYOVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSTCTHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSTCTHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1036
QY 1021 PFGIOEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1096
QY 1081 KVEPEVFNPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1137
DB 1097 KVEPEVFNPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPO 1153

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGSVVOLGSRVVVGAPQEIIVANORGSILYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGSVVOLGSRVVVGAPQEIIVANORGSILYQCDYSTGSCPEI 76

Qy 61 RLOVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLOVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 136

Qy 121 FPEALRCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVWEQLKSKTILFSLMQYSEEF 180
Db 137 FPEALRCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVWEQLKSKTILFSLMQYSEEF 196

Qy 181 RIHFTKFEQNNPRSLVPIPTQLLGRTHATGATGVRKVIKRLNITNGARKNAFKILIVI 240
Db 197 RIHFTKFEQNNPRSLVPIPTQLLGRTHATGATGVRKVIKRLNITNGARKNAFKILIVI 256

Qy 241 TDEKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTIASKPPRDHVPQVN 300
Db 257 TDEKFGDPLGYEDVIEADREGVIRVVGDAFRSEKSRQELNTIASKPPRDHVPQVN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASCLSDVDNSGSTDVLVIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASCLSDVDNSGSTDVLVIGAPHYYEOTRGQVSVCP 496

Qy 481 PRQORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRQORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIOSVVT 660
Db 617 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIOSVVT 676

Qy 661 YDLALDSGRPHSRVAVNETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

Qy 721 SLVGTPLSAFAGNLRLPVLAEDAQRLETFALFPFEKNCNDNI CQDDLSITFSFMSLDCVAVG 780
Db 737 SLVGTPLSAFAGNLRLPVLAEDAQRLETFALFPFEKNCNDNI CQDDLSITFSFMSLDCVAVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLONQSRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLONQSRQSRWLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916

Qy 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVNHQYQVSNLQORSILPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVNHQYQVSNLQORSILPISLVFLVPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSTIACQRIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSTIACQRIQCDIP 1036

Qy 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1096

Qy 1081 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMHSEGGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMHSEGGPPGAEPQ 1153

RESULT 8

US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.9%; Score 5868; DB 3; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFSGNLRRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFSGNLRRQPOK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMBOLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMBOLKSKTFLSLMQYSEEF 196

Qy 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 316

Qy 301 NFEALKTIONQREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 376

Qy 361 GVPLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVLGAPYQHIGLVAMFR 420
Db 377 GVPLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP 496

Qy 481 PRGORARWQDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGSLSGGODLTMGDLVDTVCAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGSLSGGODLTMGDLVDTVCAQGHVLLLRQ 616

Qy 601 PVLRVKAIMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVYT 660
Db 617 PVLRVKAIMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVYT 676

Qy 661 YDLALDSGRPHSAVNETKNSTRTOTVGLTQTCETLKLQPNICIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSAVNETKNSTRTOTVGLTQTCETLKLQPNICIEDPVSPVILRLNF 736

Qy 721 SLVGTPLSAGNLRPVLAEDAQRFTALPFPEKNCNDNITCODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAGNLRPVLAEDAQRFTALPFPEKNCNDNITCODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVSENMPRTNKTFF 916

Qy 901 QLELPVKYAVVMVTSHGVSSTKYNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960

Db 917 QLELPVKYAVVMVTSHGVSSTKYNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1036

Qy 1021 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1037 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1096

Qy 1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKGLFFKQYKDMMEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKGLFFKQYKDMMEGGPPGAEPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688.307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFSGNLRRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFSGNLRRQPOK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMBOLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMBOLKSKTFLSLMQYSEEF 196

Qy 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 316

Qy 301 NFEALKTIONQREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 376

```
Db 317 NFEALKTIONQREKI FAIEGTQTGSSSSPEHEMSQEGFSAAITSNQPLLSLTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYL 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYL 556
Qy 541 HGTSGSGISPSHSORIASGKSLPQYFGQSLGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLPQYFGQSLGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNPECDNQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSV 660
Db 617 PVLRVKAIMFNPREVARNPECDNQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSV 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCTEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCTEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNLRLPVLAEQAORLFTALFPFCKNCNDNICQDDLSITFFMSLDCLV 780
Db 737 SLVGTPLSAFNLRLPVLAEQAORLFTALFPFCKNCNDNICQDDLSITFFMSLDCLV 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESAS 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESAS 856
Qy 841 SGALKSTSCSINHPIFENSEVTNIITFDVDSKASLGNKLLKANVTSENNMPTNK 900
Db 857 SGALKSTSCSINHPIFENSEVTNIITFDVDSKASLGNKLLKANVTSENNMPTNK 916
Qy 901 QLELPVKYAVVMVTSYHGVSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFL 960
Db 917 QLELPVKYAVVMVTSYHGVSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFL 976
Qy 961 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVC 1020
Db 977 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVC 1036
Qy 1021 FFGIOEENATLKNLSPDFWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOT 1080
Db 1037 FFGIOEENATLKNLSPDFWIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOT 1096
Qy 1081 KVEPFEVNPPLPLIVGSSVGLLALLITAALYKLGFFKQYKDMSEGGPPGAEP 1137
Db 1097 KVEPFEVNPPLPLIVGSSVGLLALLITAALYKLGFFKQYKDMSEGGPPGAEP 1153
```

RESULT 10

```
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnsaut, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-476-062A-43
```

```
Query Match 99.4%; Score 5837.5; DB 2; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136
Qy 121 FPALAGCPQEDSDIAFLIDGSGSIIPHDPRRMEKFVSTVMEQLKSKTILFSLMOYSEEF 180
Db 137 FPALAGCPQEDSDIAFLIDGSGSIIPHDPRRMEKFVSTVMEQLKSKTILFSLMOYSEEF 196
Qy 181 RIHFTFKFONNPNRSLVKPIQTQLGRTHATGVRKVIRELNIITNGARKNAFKILIVI 240
Db 197 RIHFTFKFONNPNRSLVKPIQTQLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 300
Db 257 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQBELNTIASKPPRDHVPQVN 316
Qy 301 NFEALKTIONQREKI FAIEGTQTGSSSSPEHEMSQEGFSAAITSNQPLLSLTVGSYDWAG 360
Db 317 NFEALKTIONQREKI FAIEGTQTGSSSSPEHEMSQEGFSAAITSNQPLLSLTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYL 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYL 555
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYFGQSLGQDLTMDGLVDLTVGAQGHVLLRSQ 600
```

```
Db 556 HGTSGSGISPSHSQRIAGSKLSPLOVFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 615
Qy 601 PVLRVKAIMBPNPREVARNVFECDQVVKQKEAGEVVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 616 PVLRVKAIMBPNPREVARNVFECDQVVKQKEAGEVVRVCLHVQKSTRDRRLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 915
Qy 901 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMOHQVQVSNLQORSIPISLVFLVPV 960
Db 916 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMOHQVQVSNLQORSIPISLVFLVPV 975
Qy 961 RLNTQVWDRPQVTFSENLSSTCHTKERLPSPHSDFLAELKAPVVMVNCISIAVCQRIQCDIP 1020
Db 976 RLNTQVWDRPQVTFSENLSSTCHTKERLPSPHSDFLAELKAPVVMVNCISIAVCQRIQCDIP 1035
Qy 1021 PFGIOBEFNATLKNLSPDWIKYKSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTET 1080
Db 1036 PFGIOBEFNATLKNLSPDWIKYKSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMVSEGGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMVSEGGPPGAEPQ 1152
```

RESULT 11

```
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
```

```
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43
```

```
Query Match 99.4%; Score 5837.5; DB 5; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQSNARGFGQSVVQLOGSRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQSNARGFGQSVVQLOGSRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFRRMKFVSTVMEOLKSKTLPFLMAYSEEP 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGIIPHDFRRMKFVSTVMEOLKSKTLPFLMAYSEEP 196
Qy 181 RIHFTKFEFONNPNRSLVKPITQLLGRTHATATGVRKVIKRELLNITNGARKNAFKILVI 240
Db 197 RIHFTKFEFONNPNRSLVKPITQLLGRTHATATGIRKVVRELFNITNGARKNAFKILVI 256
Qy 241 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAFRSEKSRQBELNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAFRSEKSRQBELNTIASKPPRDHVFQVN 316
Qy 301 NFPAKTIQNLREKIFAIEGTOTGSSSEHEHMSQEGFSAAITSGNPILLSTVGSYDWAG 360
Db 317 NFPAKTIQNLREKIFAIEGTOTGSSSEHEHMSQEGFSAAITSGNPILLSTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSDMNDAYLGVAIAAILRNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSDMNDAYLGVAIAAILRNRVQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQVSVCLP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQVSVCLP 496
Qy 481 PRGARAQWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTIDVAIGAPGEDNRGAYVLF 540
Db 497 PRGARAQWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTIDVAIGAPGEDNRGAYVLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOVFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPLOVFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 615
Qy 601 PVLRVKAIMBPNPREVARNVFECDQVVKQKEAGEVVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 616 PVLRVKAIMBPNPREVARNVFECDQVVKQKEAGEVVRVCLHVQKSTRDRRLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 915
```

QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 960
DB 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 975
QY 961 RLNOQTVWDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWNCIAVCORIOCDIP 1020
DB 976 RLNOQTVWDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWNCIAVCORIOCDIP 1035
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTET 1080
DB 1036 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTET 1095
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137
DB 1096 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2
; LENGTH: 1152
5424399-2

Query Match 99.4%; Score 5837.5; DB 6; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARFGSGSVVQLQGSRVVVGAPQIIVAAQORGSILYQCDYTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGSGSVVQLQGSRVVVGAPQIIVAAQORGSILYQCDYTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVKGCLCFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVKGCLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKEFVSTVMEQLKSKTILFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKEFVSTVMEQLKSKTILFSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNSRLVPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNSRLVPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256
QY 241 TDGKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
DB 257 TDGKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 316
QY 301 NFEALKTIQOLREKIFAIEGTQTGSSSSFEHENSQSGFSAATSNGLPSTLTVGSDYDAG 360
DB 317 NFEALKTIQOLREKIFAIEGTQTGSSSSFEHENSQSGFSAATSNGLPSTLTVGSDYDAG 376
QY 361 GVFLYTSKESKSTFNNMTRVSDMNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNMTRVSDMNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMEESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGGOVSCPL 480
DB 437 QNTGMEESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGGOVSCPL 496
QY 481 PRGORARWQCDAVLYGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPEGEDNRGAVYLF 540

DB 497 PRG-RARWQCDAVLYGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPEGEDNRGAVYLF 555
QY 541 HGTSSGSI SPSHSQR IAGSKLSPRLQYFCQSLSGGQDLTMDGLVDLTGGAQGHVLLLRSQ 600
DB 556 HGTSSGSI SPSHSQR IAGSKLSPRLQYFCQSLSGGQDLTMDGLVDLTGGAQGHVLLLRSQ 615
QY 601 PVLRVKAIEMFNPREVARNVFECDQVVGKAGEVRVCLHVQKSTRDLRREQIOQSVVT 660
DB 616 PVLRVKAIEMFNPREVARNVFECDQVVGKAGEVRVCLHVQKSTRDLRREQIOQSVVT 675
QY 661 YDLALDSGRPHSRVFNETKNSRROTQVLTGTLTOTCETIKLQLPNCIEDPVSPIVLRNLF 720
DB 676 YDLALDSGRPHSRVFNETKNSRROTQVLTGTLTOTCETIKLQLPNCIEDPVSPIVLRNLF 735
QY 721 SLVGTPLSLAFGNLRPVLAEDAQRLLFTALPFFPKNCGNDNICDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSLAFGNLRPVLAEDAQRLLFTALPFFPKNCGNDNICDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLONQORSQSWRLACESASSTEV 840
DB 796 GPREFNVTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLONQORSQSWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANYTSENMPRTNKTEF 900
DB 856 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANYTSENMPRTNKTEF 915
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 960
DB 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 975
QY 961 RLNOQTVWDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWNCIAVCORIOCDIP 1020
DB 976 RLNOQTVWDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWNCIAVCORIOCDIP 1035
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTET 1080
DB 1036 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTET 1095
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137
DB 1096 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1152

RESULT 13
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 59.1%; Score 3473; DB 2; Length 1163;
Best Local Similarity 61.3%; Pred. No. 3.9e-285;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFQSVVLOGSRVVGAPQEIIVAAANQORSLYQCDYSTGSCBPI 60
DB 20 FNLDTEELTAFRVDSAGFGDSVVQVANSVVVVGAPQKITAANQOTGLYQCYGSGYACBPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSBNTVYVKGCLFGLFGLNLRQQPOK 120
DB 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRMVLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQEDSIAFLDGGSGIIPHDFRMKEFVSTVMBQLKSKTFLSLMYSEEF 180
DB 138 LPVSRQECPRQSDIVFLDGGSSISRRNFATMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RHFTTFKQNNPNRPSLVKPTQLLGRTHATGVKVIJRELLNTNGARKNAFKILIVI 240
DB 198 OTHFTPEEPRRTSNPLSLASVHLQCGFTYTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TQGEKFGDPLGVEDVTPEDREGVIRYVIGVDGAPRSEKSRQELNTIASKPRDRHVFQV 300
DB 258 TDGKEGSDLYKDVIPMDAAGIIRYAGVGLAFQNRNWSKELNDIASKPSQEHIFKVE 317
QY 301 NFEALTIQNLREKIFAIEGTQTSSEHSEMSQEGFSAAITSNGLPILLSVTGSDWAG 360
DB 318 DFDALDKIQNLKEKIFAIEGTETTSSEHSEMSQEGFSAVFTPDGVLGAVGSGFTWSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLYAAAILLRNVQSLVLGAPRYQHIGLVAMPR 420
DB 378 GAFLYPPNNSPTFINNSQENVMDRDSYLGSTELALWKGVSQSLVLGAPRYQHTGKAVIPT 437
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQVSVCP 480
DB 438 QVSRQRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVIGAPHYEQTRGGQVSVCP 497
QY 481 PRQGRMOCDAVLYGEOQPNRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVYLF 540
DB 498 PRGWR--RWWCDAVLYGEOQPNRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVYLF 556
QY 541 HTSGSGISPSHSQRTAGSKLSPRLOYFQOSLSGGQDLMGDLVLTGVAQGHVLLLRSQ 600
DB 557 HGVLGPSISPSHSQRTAGSKLSPRLOYFQOALSQGGQDLMGDLVLTGVAQGHVLLLRTR 616
QY 601 PVLRYKAIEMFNPVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLREGQOSVVT 660
DB 617 PVLWGVGVQFIPAEIPRGAFCREQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDGRPHRAVNETKNSRROTQVLGTQTCTETLQLPNCIEDPSPVILRLNF 720
DB 677 LDALDPGLSPRATQETKNSLSRVVLGLKHCENFLLPSCVEDSVPTITRLNF 736
QY 721 SLVGTPLSAFGLRPVLAEDAQRFTALFPFKNCGNDNICODDLSITFSFMSLDCLVVG 780

737 TLVGKPLLAFNRURPMLAALAQRYFTASLPFKKNCADHICQDNLGSGISFPFGLKSLLVG 796
781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWLACASSTEV 840
797 SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSYRYVAEGQKQQLRSLHLTCDSPVVG-- 854
841 SGALKSTSCSINHPIFPENSEVNTITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
855 SQGTWSTSCRINHILIFRGGAQITLFTFDVSPKAVLDGLRLLLTANVSSENNTPTTSKTF 914
901 QLELPVKYAVYVMTVSHGVSTKYLNFTAS-ENTSRVMQHOVQVSNLQORSIPLSLVLPV 959
915 QLELPVKYAVYVTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWP 974
960 VRLNQTVIWRDPQVTFSENLSSTCTTKERLPSSHDFLAELRKAPVNVNCSIAVCQRIQCDI 1019
975 VELNQAAMVDVEVSHPNPNSLRCSSEKIAPPASDFLAHQKPNVLDCSIAGCLRRCDV 1034
1020 PFGIOQEFNATLKNLSFDWYIKTSNNHLLIVSTABILFNDSVFTLLPGGAFVRSQTE 1079
1035 PSFSVQBELDFTLKNLSFGWVRQILOKQVSVSVABITFDTSVYSQLPGQAFPMRAQTT 1094
1080 TKVPEPFEVNPPLIVGSSVGLLLALITAAALYKLGFKRQYKDMHSE 1128
1095 TVLEKIVHNTPPLIVGSSVGLLLALITAAALYKLGFKRQYKDMHSE 1143

RESULT 14
PCT-US96-01314-44
Sequence 44, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaut
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44

Query Match 59.1%; Score 3473; DB 5; Length 1163;
Best Local Similarity 61.3%; Pred. No. 3.9e-285;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

Db 318 DFDALDKIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
QY 361 GVFLVTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVOSLVLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNPNSTFINNSQENVDMRDSYIGYSTELAKWGVOSLVLGAPRYOHIGKAVIFI 437
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSDVDOSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480
Db 438 QVSRQWRMKAIEVIGTIGSYFGASLCSDVDVDTGSDTDLVIGAPHYYEQTRGGQVSCPL 497
QY 481 PRGQARWQCDVAVLGEQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEBEDNRGAVYLP 540
Db 498 PRGWR-RWMCDAVLYGEOCHPMGRFGAALTVLGDVNGDKLTDVVIAPGEBEENRGAVYLP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGAGCHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRIAGSKLSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLR 616
QY 601 PVLRYKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVSNQFIPAIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVFNETHNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDLALAPGRLSPRAIFQETKRSLSRVRVLGKAHCENFNLLSPCEDSVIPIILRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAORLEFALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 TLVGPPLAFRLRPLAALAQRYFTASLPFFKNCAGADHICODNLGIGSFPGKSLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPFLDLSYRKVSTLQONORSORSWRLACESASSTEV 840
Db 797 SNLELNAEVMVWNGEDSVGTTITFSHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
QY 841 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 855 SQGTWSTSCRINHLPFGCAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
QY 901 QLELPVKYAVYVWVTSHTYKLNFTAS-ENTSRVMOHQOVSNLQORSLPISLVLVP 959
Db 915 QLELPVKYAVYVWVTSHTYKLNFTAS-ENTSRVMOHQOVSNLQORSLPISLVLVP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
Db 975 VELNQEAVMWDEVSHQPSPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDOSVFTLLPGQAFVRSQTE 1079
Db 1035 PGFSVQEEELDFTLKGNLSQWVRQILQKKVSVVVAEIIFDTSVYSQLPQOAFMRAQTI 1094
QY 1080 TKVEPPEVPNPLIIVSSVGLLALALITAAVLYKLGFEKROYKDMWSE 1128
Db 1095 TVLEKYVHNPIPLIIVSSIGGLLALITAVLYKGVGFKROYKEMWEE 1143

Search completed: November 25, 2003, 14:23:24
Job time : 16.5618 secs